New tools to verify the Collatz conjecture.

Laurent NEDELEC

E.mail : nedeleclaurent@protonmail.com

Abstract :

Since the Syracuse problem has remained unsolved for 80 years, it is possible that it is formally unsolvable. This text offers an « approximate » proof. Using the concepts of inverse graph and an original logical trick (the axis of verified integers), the Collatz conjecture is progressively verified, although no definitive conclusion is reached. Other approaches can build upon the ideas presented in this text to achieve an even more accurate approximation. The text is in French, and an English version is expected to be completed within two months.

Présentation du problème

Le problème de Collatz (ou problème (3n + 1), ou problème de Syracuse) a été proposé pour la première fois par Lothar Collatz dans les années 1930. Son énoncé est extrêmement simple : l'application de Collatz associe à tout entier positif n un entier unique n' tel que n' = n/2 si n est pair, et n' = (3n + 1)/2 si n est impair. Les résultats des itérations successives de l'application de Collatz à partir d'un entier n quelconque sont appelés "trajectoire de Syracuse de n".

Le problème de Collatz consiste à démontrer si oui ou non toutes les trajectoires de Syracuse, à partir de n'importe quel entier positif, atteignent la valeur 1 en un nombre fini d'itérations. La conjecture de Collatz est l'hypothèse que la solution au problème de Collatz est positive, c'est à dire que toutes les trajectoires de Syracuse tendent vers 1.

Malgré son énoncé très simple ce problème est irrésolu depuis plus de 80 ans. La difficulté principale vient du fait que les opérations de multiplication par (3n + 1)/2 ou de division par 2 selon la parité des entiers obtenus lors des itérations successives de l'algorithme, sont difficiles à prévoir en fonction de la valeur de départ, et produisent des résultats imprévisibles, se rapprochant de phénomènes aléatoires.

L'approche proposée dans ce texte est de nature algorithmique et propose, grâce à divers procédés logiques (notamment le concept d'ensemble des vérifiés) et à une utilisation des notions de graphe inverse, de donner une réponse positive au problème de Collatz. L'intérêt de ce texte est que certaines démonstrations, obtenues par d'autres moyens, peuvent éventuellement se rajouter aux diverses techniques proposées, et confirmer le résultat du texte par d'autres procédés.

PARTIE I

Définitions liées au problème de Collatz

1) L'application de Collatz ou Syracuse

Définissons l'application C, dite de Collatz, ou de Syracuse, qui associe à tout entier n € N* un unique entier C(n) € N* tel que :

si *n* est pair, C(n) = n / 2

si *n* est impair, C(n) = (3n + 1) / 2

2) Les trajectoires T_n

Quels que soient n € N* et i € N, nous allons définir une suite T_n (trajectoire de Syracuse ayant pour valeur de départ *n*) composée d'entiers que l'on représentera sous la forme S_i (n) € N* avec :

S_0 (n) = n

S_{i+1} (n) = C(S_i (n)) C étant l'application définie précédemment.

Nous pouvons donc exprimer T_n de la manière suivante : quels que soient *n* et *i* appartenant à N*

T_n = { n ; S_1 (n) ; ... ; S_i (n) ; S_{i+1} (n) ; ...}

Nous avons par exemple :

T_1 = { 1; 2 ; 1 ; 2 ; 1 ; ...}

T_3 = { 3 ; 5 ; 8 ; 4 ; 2 ; 1 ; 2 ; ... }

3) L'ensemble des entiers vérifiés V

Nous définirons *n* comme étant « un entier vérifié », s'il existe un entier *k* tel que S_k (n) = 1

Nous appellerons V l'ensemble des entiers vérifiés.

4) Vérification d'une trajectoire

La vérification d'une trajectoire T_n consiste à déterminer si cette trajectoire converge vers 1, c'est à dire s'il existe un entier *k* tel que $S_k(n) = 1$.

5) Itérations paires et impaires

Dans la suite du texte, les termes généraux « itération impaire » ou « itération paire » désigneront une valeur, respectivement impaire ou paire, d'un $S_k(n)$ quelconque appartenant à un T_n quelconque (amenant donc une itération impaire de la forme $C(S_k(n)) = (3S_k(n) + 1) / 2$ ou une itération paire de la forme $C(S_k(n)) = S_k(n) / 2$ dans la trajectoire T_n, suivant la parité de $S_k(n)$). Dire « il y a trois itérations paires successives dans cette trajectoire », signifie que trois valeurs paires se sont succédées, entraînant trois itérations de type « pair » dans cette trajectoire de Syracuse.

PARTIE II

Règles d'organisation de l'ensemble des vérifiés

Nous pouvons initialiser l'ensemble des vérifiés V en vérifiant successivement les trajectoires pour les premiers entiers à partir de 1 :

T_1 = { 1 ; 2 ; 1 ; 2 ... }
T_2 = { 2 ; 1 ; 2 ; 1 ... }
T_3 = { 3 ; 5 ; 8 ; 4 ; 2 ; 1 ; 2 ... }
T_5 = { 5 ; 8 ; 4 ; 2 ; 1 ; 2 ... }
T_7 = { 7 ; 11 ; 17 ; 26 ; 13 ; 20 ; 10 ; 5 ; 8 ; 4 ... }
T_9 = { 9 ; 14 ; 7 ; 11 ; 17 ... }
T_{11} = { 11 ; 17 ; 26 ; 13 ... }

Pour les 11 premières trajectoires, on constate que l'ensemble des vérifiés est égal à : (1 ; 2 ; 3 ; 4 ; 5; 6 ; 7 ; 8 ; 9 ; 10 ; 11 ; 13 ; 14 ; 17). On constate que les trajectoires 4 ; 6 ; 8 et 10 n'ont pas besoin d'être étudiées, car leur valeur initiale étant paire, elles sont dès la première itération divisées par 2, et rejoignent la trajectoire d'une valeur vérifiée précédemment. De plus T_5 n'avait pas besoin d'être calculée également car la valeur 5 est apparue dans T_3 . Nous savons donc l'évolution que donnera la valeur 5, il suffit de voir comment a évolué T_3 .

Nous allons formaliser ces quelques remarques. En effet l'ensemble des vérifiés comporte des règles internes :

1) Toute valeur d'une trajectoire vérifiée est automatiquement considérée comme vérifiée

Nous avons :

$T_1 = \{ 1 ; 2 ; 1 ; 2 ; 1 ; ...\}$

T_1 est appelé le cycle trivial, car dès qu'une trajectoire est égale à la valeur 1, alors elle rentre dans ce cycle trivial, alternance de 1 et de 2.

Supposons que n € V. Cela veut dire qu'il existe un entier *k* tel que $S_k (n) = 1$.

Nous avons donc $T_n = \{ n ; S_1 (n) ; \dots ; S_{k-1} (n) \} + T_1$

Tous les $S_p (n)$ avec p € [1 ; k-1] appartiennent à V

En effet, comme tous les $S_p (n)$ sont des entiers uniques, clairement déterminés et ordonnés par l'algorithme de Collatz, nous avons :
$T_{Sp(n)} = \{ S_p (n) ; S_{p+1} (n) ; \dots ; S_{k-1} (n) \} + T_1$

Par conséquent nous obtenons le théorème suivant :

THEOREME 1

Si n € V cela implique qu'il existe un entier *k* tel que S_k (n) = 1. L'ensemble des entiers S_p (n), quel que soit p € [1 ; k-1], appartient à V.

Note : pour prendre un exemple,

T_3 = { 3 ; 5 ; 8 ; 4 ; 2 ; 1 ; 2 ; ... }.

L'entier impair 5 apparaît dans T_3 qui atteint 1 en un nombre fini d'itérations. Donc 5 est un entier vérifié, on sait que T_5 atteint 1 en un nombre fini d'itérations, il n'y a donc pas besoin de calculer les termes successifs de T_5 pour obtenir la certitude que T_5 est vérifiée.

2) Quand une trajectoire non-vérifiée est égale à un entier vérifié, elle est automatiquement vérifiée

Quels que soient n € N* et j € N, si S_j (n) = n' avec n' € V, alors il existe un entier fini *m* tel que :

$T_{n'}$ = { n' ; S_1 (n') ; ... ; S_{m-1} (n') } + T_1

D'où

T_n = { n ; S_1 (n) ; ... ; S_{j-1} (n) } + { n' ; S_1 (n') ; ... ; S_{m-1} (n') } + T_1

j et *m* sont des entiers, donc T_n atteint l'entier 1 en un nombre fini d'itérations. *n* est donc un entier vérifié.

Par conséquent nous obtenons le théorème suivant :

THEOREME 2

Quels que soient n € N* et j € N*, si S_j (n) = n' avec n' € V, alors n € V.

Note : pour prendre un exemple, T_9 = { 9 ; 14 ; 7 ; 11 ; 17 ... }. 7 est un entier vérifié car T_7 est vérifié. T_9 atteint un entier vérifié en un nombre fini d'itérations. 9 est donc un entier vérifié : nous n'avons pas à poursuivre les calculs jusqu'à atteindre 1 pour obtenir cette certitude.

3) Chaque entier vérifié multiplié par l'ensemble des 2^d est un entier vérifié

Supposons n € V. Cela implique qu'il existe un entier *k* tel que S_k (n) = 1.

Quel que soit *d* un entier appartenant à N*, on a :

$T_{n \cdot 2^d} = \{ n \cdot 2^d ; n \cdot 2^{d-1} ; ... ; 2n \} + \{ n ; S_1 (n) ; \dots ; S_{k-1} (n) \} + T_1$

k et *d* sont des entiers d'où $T_{n \cdot 2^d}$ atteint en un nombre fini d'itérations l'entier 1.

Par conséquent, nous obtenons le théorème suivant :

THEOREME 3

Quels que soient n et d € N*, si n € V , alors $n \cdot 2^d$ € V.

4) Décomposition de l'ensemble des vérifiés

Nous allons décomposer l'ensemble des vérifiés V en trois sous-ensembles :

a) un sous-ensemble fini A(n) constitué d'un bloc de départ [1 ; n] d'entiers vérifiés (les calculs sur ordinateur ont par exemple permis de montrer que tous les entiers inférieurs à $n = 10^{20}$ vérifient la conjecture de Syracuse).

b) un sous-ensemble fini B(n) d'entiers impairs vérifiés appartenant à [n+1 ; +∞[(B(n) est constitué d'entiers impairs vérifiés, qui sont supérieurs à n, et qui sont apparus lors des vérifications successives des trajectoires de A(n), voir le **théorème 1**).

c) un sous-ensemble infini C(n) d'entiers pairs vérifiés appartenant à [n+1 ; +∞[(C(n) est constitué de l'ensemble des entiers vérifiés appartenant à A(n) et B(n) multipliés par l'ensemble des 2^d , d € N*, voir le **théorème 3**).

Nous pouvons remarquer que V est un ensemble infini. De plus, chaque nouvelle trajectoire vérifiée génère une infinité de nouveaux entiers vérifiés.

PARTIE III

L'axe des vérifiés

Considérons l'ensemble N*, comme un axe vertical, partant de 1 et allant jusqu'à l'infini, constitué d'un empilement de carrés égaux, chaque carré, ou case, étant caractérisé par sa valeur numérique donnant son positionnement dans N*.

Nous allons appeler « axe des vérifiés » un tel axe, défini de la manière suivante :

La case pour une valeur numérique donnée est blanche si la valeur en question n'appartient pas à V. La case en question est dite non-vérifiée.

La case pour une valeur numérique donnée est noire si la valeur en question appartient à V. La case en question est dite vérifiée.

Une case non-vérifiée peut devenir vérifiée.

Une fois qu'une case est vérifiée, elle le reste en permanence.

Si une trajectoire lors de ses itérations est égale à une case vérifiée, alors toutes les cases de la trajectoire sont vérifiées et deviennent noires (**théorèmes 1 et 2**). L'ensemble des cases qui ont une valeur égale à 2^d multipliée par la valeur d'une case vérifiée sont vérifiées, pour tout entier *d*. (**théorème 3**)

En fonction de la parité des résultats des itérations, les suites montent (grâce à une ou des itérations Impaires successives) ou descendent (grâce à une ou des itérations Paires successives) le long de l'axe des vérifiés.

Au fur et à mesure des vérifications successives des T_n, le nombre d'entiers vérifiés le long de l'axe augmente (**théorème 1**).

Considérons cet axe comme doté d'un ensemble de départ A(n). Puisque cet axe a un ensemble de départ A(n), il est également doté d'un ensemble fini B(n) et d'un ensemble infini C(n) (voir le point **II 4**)).

PARTIE IV

Les trajectoires majorées et sans cycles non-triviaux sont vérifiées.

Prenons l'hypothèse dans un premier temps que les trajectoires de Syracuse que nous analysons sont toutes majorées et qu'elles n'ont pas de cycles non-triviaux. Le problème de l'axe des vérifiés se transforme alors en un problème de crible, de nature algorithmique.

Le cas des cycles non-triviaux sera abordé par la suite (**Partie VI**).

Le cas des trajectoires non majorées sera également analysé par la suite (**partie VII**) : nous prendrons pour hypothèse qu'il existe une trajectoire divergente, donc, non-majorée, et nous étudierons les probabilités d'existence d'une telle trajectoire.

1) Le majorant M

Supposons que l'axe des vérifiés est doté d'un ensemble de départ A(n). Considérons la trajectoire ayant pour valeur de départ *n*', entier non-vérifié. Nous supposons que la trajectoire de Syracuse de n' est majorée par M.

2) L'intervalle [n,M] contient des entiers vérifiés

Sur l'intervalle [n ; 2n], tous les entiers pairs sont vérifiés, car tous les entiers de A(n) sont vérifiés.

Sur l'intervalle [2n ; 4n], la moitié des entiers pairs sont vérifiés.

Sur l'intervalle $[n\cdot 2^{i-1} ; n\cdot 2^{i}]$, une proportion de $1/2^{i-1}$ de tous les entiers pairs sont vérifiés.

Ainsi l'intervalle [n ; M] contient au moins ces entiers vérifiés, tout du moins pour ceux inférieurs à M.

3) Utilisation du crible des vérifiés

On prend comme hypothèse qu'il n'y a pas de cycle non-trivial dans la trajectoire considérée. Ainsi la trajectoire de *n*' évolue entre 3 et M, sans rentrer dans un cycle.

Par définition de l'application de Collatz, $C(n) \neq n$ pour tout entier n € N*. Comme la trajectoire est majorée par M et que nous avons supposé qu'elle n'a pas de cycle non-trivial, la trajectoire va prendre des valeurs à chaque fois différentes dans l'intervalle [3 ; M].

Puisqu'il y a un nombre non nul d'entiers vérifiés dans l'intervalle [n ; M], la trajectoire de Syracuse de n' n'a alors que trois choix :

a) Soit elle devient égale à un entier vérifié de [n ; M] et est donc vérifiée.

b) Soit, elle descend en dessous de *n*, devient égale à une valeur de A(n), et est donc vérifiée.

c) Soit elle continue ses itérations, jusqu'à épuiser les possibilités d'entiers non-vérifiés contenus dans [n ; M]. En effet, la trajectoire ne peut pas être égale à une constante, car $C(n) \neq n$, et elle ne peut pas être égale deux fois à la même valeur, sinon, nous serions en présence d'un cycle non-trivial, car chaque trajectoire emprunte un chemin clairement défini qui, s'il revient sur une valeur déjà prise, rentrera forcément dans un cycle non-trivial.

Quand toutes les possibilités d'entiers non-vérifiés contenus dans [n ; M] sont épuisées, l'étape suivante pour la trajectoire consiste à être égale à un entier vérifié et donc à être vérifiée.

Nous arrivons donc au résultat suivant :

THEOREME 4

Considérons un axe des vérifiés avec un ensemble de départ vérifié A(n). Si la trajectoire de Syracuse de *n'*, entier non-vérifié quelconque, est majorée, et qu'elle n'a pas de cycle non-trivial, alors elle est vérifiée.

Nous pouvons déduire de ce théorème que tant qu'il n'y a pas de trajectoire non-majorée, c'est à dire divergente, et tant qu'il n'y a pas de cycle non-trivial, les trajectoires avec des valeurs de départ supérieures à A(n) sont successivement et automatiquement vérifiées.

Les seuls problèmes qui peuvent empêcher ce processus de vérification sont donc la possibilité de trajectoires divergentes et de trajectoires avec des cycles non-triviaux. Nous allons examiner les probabilités d'existence de telles trajectoires (**parties VI et VII**). Pour cela nous allons d'abord, dans la **partie V**, aborder la question des graphes inverses.

PARTIE V

Les graphes inverses

1) Présentation du problème

Nous pouvons constater que l'application de Collatz n'est pas injective car, pour ne prendre qu'un exemple, C(3) = C(10) = 5. Cela veut donc dire que certaines valeurs peuvent avoir plusieurs antécédents.

Définition : si nous appelons *z* l'antécédent de *n* par l'application de Collatz, nous avons C(z) = n.

Ainsi 3 et 10 sont des antécédents de 5. Tout entier *n* admet au moins un antécédent, venant d'une itération paire, c'est le double de l'entier en question. En effet C(2n) = n. Par contre, selon certaines conditions que nous allons analyser, l'entier *n* peut avoir un antécédent d'une valeur inférieure à *n*, venant d'une itération Impaire.

Prenons l'exemple de l'entier vérifié 5 :

A partir du 5, en multipliant par les puissances de 2, le 10, le 20, le 40 et le 80 sont par exemple automatiquement vérifiés. Et certaines de ces valeurs ont des antécédents impairs :
on a par exemple C(3) = 5 ; C(13) = 20 ; C(53) = 80. Puisque les entiers 3, 13 et 53 amènent par l'application de Collatz à un entier vérifié, ils sont vérifiés (**théorème 2**). Eux-même auront des antécédents, et ainsi de suite. Le 10 et le 40 n'ont pas d'antécédents impairs, ils n'ont que des antécédents pairs.

Nous voyons donc qu'à partir de l'entier 5, des antécédents sont connus : 3, 10, 20, 40, 80. Ces antécédents ont eux-même des antécédents : 13, 53... Nous pouvons donc, à partir d'un entier quelconque *n*, créer une infinité d'antécédents, car nous aurons au moins l'infinité des $n \cdot 2^d$ comme antécédents.

2) La construction des graphes inverses

Nous allons présenter plus en détail ce que nous pouvons appeler le graphe inverse de Collatz :

Supposons que nous partions d'un entier *n,* et que nous cherchions à savoir si *n* a un antécédent impair m = 2k +1 tel que (3m + 1)/2 = n.

On obtient facilement m = (2n – 1)/3.

D'où 2k + 1 = (2n – 1)/3 et n = 3k + 2

Donc, si n = 3k +2, *n* admet un antécédent impair de la forme m = 2k + 1.

Si n = 3k, 2n – 1 = 6k – 1 qui n'est jamais divisible par 3.

Si $n = 3k + 1$, $2n – 1 = 6k + 1$ qui n'est jamais divisible par 3.

Donc, si $n = 3k$ ou $n = 3k + 1$, *n* n'admet pas d'antécédents impairs.

Nous pouvons alors définir la relation inverse de Collatz de *n*, qui donne les antécédents de *n*, et que nous appellerons I :

Définition : Nous pouvons définir la relation de Collatz inverse I qui associe à tout *n*, les entiers I(n), les antécédents de *n*, définis par :

$I(n) = \{2n\}$ si $n = 3k$ ou $n = 3k + 1$

$I(n) = \{2n ; (2n – 1)/3\}$ si $n = 3k +2$

Nous voyons donc que I(n) nous donne les antécédents de *n* par l'application de Collatz. A partir de la relation I, nous pouvons constituer par itérations successives, le graphe inverse de Collatz d'un entier *n* quelconque, que nous appellerons G(n).

Définition : A partir d'un entier *n* quelconque, nous recherchons I(n), c'est à dire les antécédents de *n*. A la première étape, les éléments du graphe inverse $G_1(n)$ sont alors égaux à I(n).

A la deuxième étape, nous recherchons les antécédents des antécédents que nous avons obtenus à la première étape. A la deuxième étape $G_2(n)$ est égal à $G_1(n)$ auquel on ajoute les nouveaux antécédents que l'on vient de trouver.

A la *(k+1)* étape, nous recherchons les antécédents des antécédents trouvés à l'étape *k* . $G_{k+1}(n)$ est égal à $G_k(n)$ auquel on ajoute les antécédents que l'on vient de trouver à l'étape *(k+1)*.

En reproduisant le processus d'une manière infinie, nous obtenons G(n), le graphe inverse de *n.*

Cela veut dire que toute trajectoire de Collatz, ayant comme point de départ tout entier appartenant à G(n), graphe inverse de *n*, amènera forcément à *n*.

Nous pouvons détailler ce processus de cette manière :

Supposons que $I(n) = \{n_1 ; n_2\}$

On a alors $G_1(n) = \{n_1 ; n_2\}$

Supposons que $I(n_1) = \{n_3\}$ et $I(n_2) = \{n_4 ; n_5\}$

On a alors $G_2(n) = G_1(n) \cup \{ n_3 ; n_4 ; n_5\}$

Supposons que $I(n_3) = \{n_6 ; n_7\}$, $I(n_4) = \{n_8\}$ et $I(n_5) = \{n_9 ; n_{10}\}$

On a alors $G_3(n) = G_2(n) \cup \{n_6 ; n_7 ; n_8 ; n_9 ; n_{10}\}$

Et ainsi de suite.

Prenons l'exemple de n = 17 :

17 est de la forme 3k + 2. Donc I(17) = {34 ; 11}.

On analyse ensuite I(34) = {68}   I(11) = {22 ; 7}

Puis I(68) = {136 ; 45}   I(22) = {44}   I(7) = {14}

Nous voyons donc que le graphe inverse G(17) est successivement égal à :

$G_1(17) = \{11 ; 34\}$

$G_2(17) = \{7 ; 11 ; 22 ; 34 ; 68\}$

$G_3(17) = \{7 ; 11 ; 14 ; 22 ; 34 ; 44 ; 45 ; 68 ; 136\}$

Toutes les valeurs de $G_3(17)$ ont des trajectoires qui mènent à 17.

$S_3\,(136) = 17$

$S_3\,(44) = 17$

$S_2\,(7) = 17$

3) Premiers résultats

a) Tout d'abord, le graphe inverse d'un entier *n* inclut les graphes inverses de tous les éléments qui constituent le graphe inverse de *n*. En effet, on calcule progressivement tous les antécédents de *n*, puis les antécédents de ces antécédents, ceci d'une manière infinie. Les antécédents de chacun de ces antécédents, c'est à dire le graphe inverse de chacun de ces antécédents, seront inclus dans le graphe inverse de l'entier *n* de départ.

Si $G(n) = \{n_1 ; n_2 ; \ldots ; n_i ; n_{i+1} ; \ldots\}$ alors nous avons

$G(n) = G(n_1) \cup G(n_2) \cup \ldots \cup G(n_i) \cup G(n_{i+1}) \cup \ldots$

Nous obtenons donc le résultat suivant :

THEOREME 6

Le graphe inverse d'un entier *n* inclut tous les graphes inverses des éléments qui composent le graphe inverse de *n*.

b) D'autre part, nous pouvons remarquer que le graphe inverse de n'importe quel *n* est un ensemble infini. En effet, selon la construction des graphes inverses, chaque graphe inverse G(n) comporte au moins une infinité de valeurs correspondant à l'ensemble des 2^d , quel que soit *d*, multipliés par *n*.

Nous obtenons donc le résultat suivant :

THEOREME 7

Le graphe inverse de n'importe quel entier *n* est un ensemble infini.

c) De plus, supposons que nous cherchions le graphe inverse d'un entier vérifié *n*. Puisque chaque trajectoire de Syracuse ayant un élément du graphe inverse comme point de départ, amène à *n*, toutes ces trajectoires sont vérifiées (voir le **théorème 1 et 2**). Nous pouvons donc dire que si *n* est vérifié, tout le graphe inverse de *n* est vérifié.

Ainsi, nous obtenons le résultat suivant :

THEOREME 8

Tous les éléments du graphe inverse d'un entier vérifié sont vérifiés.

d) Enfin, nous pouvons dire que si un seul élément d'un graphe inverse d'un entier quelconque est vérifié, alors tout le graphe est vérifié.

En effet, toutes les trajectoires ayant une valeur de départ du graphe inverse d'un entier quelconque *n* mènent à cet entier *n*. Si une des valeurs de ce graphe inverse est égale à un entier vérifié, cela veut dire que sa trajectoire mène à 1. Ainsi, la trajectoire égale à un entier vérifié mène à *n*, mais aussi à 1. Cela veut donc dire que *n* converge vers 1, et que donc *n* est vérifié. Comme *n* est vérifié, tout son graphe inverse est vérifié (**théorème 8**)

Nous arrivons donc au résultat suivant :

THEOREME 9

Si un élément d'un graphe inverse d'un entier quelconque est vérifié, alors tout le graphe inverse est vérifié.

4) Application des résultats à G(1)

Nous pouvons remarquer que tous les graphes inverses des entiers vérifiés sont inclus dans le graphe inverse de 1, c'est à dire G(1), car, par définition, toutes les valeurs de ces graphes inverses ont des trajectoires qui mènent à 1.

Les graphes inverses des différents entiers vérifiés s'ajoutent sur l'axe des vérifiés et viennent le noircir : chaque nouvel entier vérifié diminue le nombre de cases blanches, c'est à dire non-vérifiées. Cela se fait d'une manière automatique, au fur et à mesure des itérations permettant de découvrir, avec l'algorithme décrit précédemment, de nouveaux éléments des graphes inverses des entiers vérifiés.

La question qui se pose est la suivante : est-ce que la réunion de tous les graphes inverses de tous les entiers vérifiés, c'est à dire G(1), couvre N* ? Cela voudrait dire que G(1) remplit progressivement l'axe des vérifiés, ne laissant plus aucun entier non-vérifié. La conjecture de Collatz peut d'ailleurs se reformuler de la manière suivante : prouver que le graphe inverse G(1) est égal à N*, ce qui équivaut à dire que toutes les trajectoires convergent vers 1.

Les vérifications expérimentales sur ordinateur prouvent que toutes les trajectoires des entiers jusqu'à 10^{20} convergent, sans exception, vers 1. Cela veut dire que tous ces entiers jusqu'à 10^{20} sont contenus dans G(1). Cela veut donc aussi dire que G(1) est au moins la réunion de tous les graphes inverses des entiers jusqu'à 10^{20} sans exception. Comme un graphe inverse contient une infinité de valeurs, on comprend que la réunion de tous ces graphes inverses va remplir l'axe des vérifiés en vérifiés, et que cet axe va s'obscurcir progressivement, et d'une manière irréversible. G(1) est déjà, selon l'avancement de nos connaissances, extrêmement grand, puisqu'il couvre complètement tous les entiers jusqu'à 10^{20}, et inclut tous leurs graphes inverses.

Les temps de vol maximum pour les entiers jusqu'à 10^{20} ne dépassent pas 3000, cela veut dire qu'il faut un maximum de 3000 itérations pour une trajectoire de Syracuse pour aller d'un point de départ quelconque dans [1 ; 10^{20}] jusqu'à 1 (voir sur le site internet de Eric Rosendaal : ericr.nl « 3x+1 delay records »). Cela veut donc aussi dire, dans le sens inverse, qu'un maximum de 3000 itérations du graphe inverse de 1 suffisent à couvrir entièrement l'intervalle [1 ; 10^{20}] en vérifiés.

Cependant le graphe $G_{3000}(1)$ s'étend, sans aucun doute, bien au-delà de la limite de [1 ; 10^{20}], où tous les entiers sont vérifiés d'une manière sûre, et que l'on pourrait identifier à l'ensemble de départ A(n) (voir **partie II point 4)**). Au delà de cette limite se mélangent des entiers appartenant à $G_{3000}(1)$ et des entiers pas encore vérifiés. Ce mélange de vérifiés et de non-vérifiés est présent depuis la borne supérieure de A(n), qui est de 10^{20} environ pour $G_{3000}(1)$, comme nous venons de le voir, jusqu'à 2^{3000} qui est la borne supérieure de $G_{3000}(1)$.

On peut noter qu'aux alentours de 2^{3000} il y aura peu d'éléments de $G_{3000}(1)$. Au contraire, si on appelle n' le premier entier non vérifié au dessus de la borne supérieure de A(n), il est très probable que cet entier n' soit entouré de beaucoup de vérifiés et que l'ensemble des non-vérifiés, c'est à dire l'ensemble des cases non-vérifiées restantes dans les alentours de n', soit extrêmement discontinu. Plus A(n) deviendra grand et plus l'ensemble des entiers non-vérifiés au dessus de la borne supérieure de A(n) risquera de devenir discontinu. Les valeurs non-vérifiées restantes seront en quelque sorte les derniers éléments attendant d'être vérifiés lors des prochaines itérations de G(1), cet ensemble devenant de plus en plus énorme et remplissant progressivement l'axe des vérifiés.

Cette forte concentration en entiers vérifiés dans la proximité de la borne supérieure de A(n) risque de rendre difficile les débuts d'une éventuelle trajectoire non-vérifiée. En effet, une trajectoire non-vérifiée doit aller d'entiers non-vérifiés à entiers non-vérifiés sans jamais être égale à un entier vérifié. Dans ce contexte, les premières itérations de n', le premier entier non-vérifié au dessus de A(n), risquent d'être rapidement égales à un entier vérifié.

Quand on voit que 3000 itérations de construction de G(1) suffisent à vérifier l'ensemble des entiers sur des centaines de milliards de milliards d'entiers, on peut se demander quelle sera la borne supérieure de A(n) lorsque des milliards et des milliards d'itérations construisant chaque graphe inverse de chaque entier vérifié seront appliquées : il est très probable que plus le nombre d'itérations appliquées à G(1) augmentera, et plus l'ensemble de départ A(n) couvrira progressivement des intervalles de plus en plus grands de N*.

Nous devons noter que la construction des graphes se fait d'une manière algorithmique. Cette approche est donc dépendante du temps. Cependant, nous pouvons considérer l'ensemble du graphe inverse de

1, ou de n'importe quel autre entier, comme déjà existant : l'ensemble de ces valeurs amenant à 1 sont, à un niveau philosophique, existantes, déjà présentes, vraies, réelles, même si nous n'avons pas la confirmation algorithmique et numérique exacte de toutes ces valeurs au moment présent. L'approche algorithmique ne fait que révéler pas à pas l'étendue infinie et réelle de chaque graphe inverse, et notamment celui de G(1).

5) Hypothèse que G(1) ne couvre pas tout N*

Nous avons vu que la réunion de milliards de milliards de graphes inverses d'entiers déjà vérifiés rend l'axe des vérifiés de plus en plus sombre. Ceci laisse à penser que G(1) couvre tout N*.

Supposons que G(1) ne couvre pas tout N*. Cela signifierait qu'il existe des entiers dont la trajectoire ne converge pas vers 1. Cela voudrait dire que les entiers qui ne rentrent pas dans G(1) font partie d'une trajectoire divergente, ou d'un cycle non-trivial (en effet ce sont les seuls cas qui restent, car une trajectoire majorée et sans cycle non-trivial est automatiquement vérifiée, voir la **partie IV**).

Quel est, alors, le sens des cases blanches restant sur l'axe des vérifiés, au dessus de A(n), symbolisant les hypothétiques derniers entiers non-vérifiés ayant échappé à la réunion de plus en plus énorme des milliards de milliards de graphes inverses d'entiers vérifiés ? Il faut avouer que ces cases blanches sur l'axe des vérifiés, de plus en plus rares et de plus en plus discontinues, ont de moins en moins de sens. En effet, comment rejoindre entre eux ces entiers non-vérifiés de plus en plus isolés ? Toute trajectoire partant de la valeur de départ n', premier entier non-vérifié supérieur à A(n), aura de très grandes probabilités de rencontrer un entier vérifié avant de rejoindre un autre entier non-vérifié. En effet, au fur et à mesure des itérations construisant les graphes inverses de tous les entiers vérifiés, l'ensemble des non-vérifiés devient de plus en plus discontinu. Plus le nombre d'itérations menant d'entiers non-vérifiés à entiers non-vérifiés doit être long, et plus les probabilités de l'existence d'une telle trajectoire vont en diminuant. Peu de solutions sont proposées aux entiers non-vérifiés, à part se transformer en entier vérifié.

Plus nous vérifions d'entiers grâce aux itérations de G(1), et plus les probabilités pour les trajectoires de Syracuse restantes de rencontrer un entier vérifié augmentent, ce qui les amènent à être de plus en plus rapidement vérifiées, et à ainsi créer encore plus de vérifiés. En effet, toutes ces trajectoires vérifiées et tous ces graphes inverses n'évoluent que sur un seul et même axe, l'axe des vérifiés, qui, au fur et à mesure de chaque vérification de trajectoire ou d'avancée de G(1), se remplit en entiers vérifiés.

Nous pourrions résumer cela de la manière suivante : les dernières hypothétiques cases blanches, symbolisant les entiers non-vérifiés et restant à vérifier, se vérifient en quelque sorte d'elles-mêmes, car ces entiers non-vérifiés qui restent sont de plus en plus entourés d'entiers vérifiés : les trajectoires qui auraient pu « soutenir » l'évolution de ces valeurs non-vérifiées vers une divergence, ou un cycle non-trivial (en effet, ce sont les seuls cas qui restent, voir la **partie IV** et le **théorème 4**), ne peuvent pas se déployer, sans rapidement être égales à un entier vérifié, car ceux-ci sont omniprésents (comme le montrent les vérifications expérimentales couvrant tous les entiers jusqu'à 10^{20} pour $G_{3000}(1)$). Ainsi, les dernières trajectoires comportant des entiers non-encore vérifiés se vérifient d'une manière de plus en plus rapide, quasi-automatiquement, car il y a un manque de non-vérifiés. Certaines approches « probabilistes » du problème de Syracuse, amènent à la conclusion que presque toutes les trajectoires convergent vers 1 (voir par exemple T.Tao : *Almost all orbits of the Collatz map attain almost bounded values*). Dans l'optique de l'axe des vérifiés, de telles trajectoires avec des possibilités de divergence, ou comportant des risques de cycle non-trivial, et qui ont des probabilités quasiment nulles d'exister, n'auraient en quelque sorte pas assez « d'espaces libres », de cases blanches, c'est à dire d'entiers non-vérifiés, pour se « déployer ». Cela va être le thème des prochaines **parties VI et VII**.

PARTIE VI

La question des cycles non-triviaux

Dans la **partie IV** nous avons montré que toutes les trajectoires majorées et qui n'ont pas de cycles non-triviaux, sont vérifiées. Nous allons maintenant examiner l'hypothèse qu'une des trajectoires puisse avoir un cycle non-trivial. Si nous prenons comme hypothèse qu'il existe un cycle non-trivial dans une des trajectoires, alors, cette trajectoire est majorée, car la trajectoire ne peut pas à la fois rentrer dans un cycle non-trivial et diverger. Les deux propositions sont contradictoires. En effet, dès qu'une trajectoire rentre dans un cycle non-trivial, elle continue à répéter ce cycle d'une manière infinie. La trajectoire est majorée par la valeur maximale M du cycle non-trivial, ou par la valeur maximale M' prise par la trajectoire avant de rentrer dans le cycle non-trivial. Dans les deux cas la trajectoire ayant un cycle non-trivial est majorée et évolue sur un axe des vérifiés de plus en plus rempli en vérifiés.

On peut remarquer également que s'il existe une trajectoire avec un cycle non-trivial, alors il existe une infinité de trajectoires avec un cycle non-trivial. En effet, considérons le graphe inverse de n'importe quelle valeur de la trajectoire ayant un cycle non-trivial. Toutes les valeurs de ce graphe inverse auront des trajectoires qui mèneront à la trajectoire ayant un cycle non-trivial. De plus, le graphe inverse de n'importe quel entier est infini (**théorème 7**).

Nous obtenons donc le résultat suivant :

THEOREME 10

S'il existe une trajectoire avec un cycle non-trivial, alors il existe une infinité de trajectoires avec un cycle non-trivial.

Aucun élément du graphe inverse des valeurs du cycle non-trivial ne peut avoir de valeurs égales à un entier vérifié (**théorème 9**) car si une seule de ces valeurs est égale à un entier vérifié, cela veut dire qu'au moins un terme du cycle non-trivial converge vers 1, ce qui serait contradictoire avec le fait que la trajectoire a un cycle non-trivial.

Nous obtenons donc le résultat suivant :

THEOREME 11

Si une seule des valeurs de la réunion des graphes inverses de tous les termes d'un hypothétique cycle non-trivial est égale à un vérifié, alors le cycle ne peut pas exister et toutes ces valeurs convergent vers 1.

En plus du **théorème 11**, les caractéristiques d'un cycle non-trivial dans une trajectoire de Syracuse font que ceux-ci sont soumis à diverses contraintes. Une analyse effectuée par S. Eliahou sur la structure des cycles non-triviaux dans les trajectoires de Syracuse (S. Eliahou : *The 3x+1 problem : new lower bounds on non-trivial cycle lengths*) affirme que si les 2^{52} premiers entiers sont vérifiés (ce qui est le cas, puisque les 2^{68} premiers entiers sont vérifiés) alors la borne inférieure pour le nombre d'itérations d'un cycle non-trivial sera de 187363077. Nous savons également que plus l'ensemble de départ des entiers vérifiés A(n) augmente, plus la borne inférieure pour le nombre d'itérations d'un cycle non-trivial augmente. Si la borne supérieure de A(n) est supérieure à 2^{1000}, alors la borne inférieure pour le cycle non-trivial sera approximativement de $3{,}45 \cdot 10^{500}$.

Nous sommes donc en face d'une double contrainte : plus l'ensemble croissant que constitue la réunion des multiples graphes inverses de tous les entiers déjà vérifiés (au fur et à mesure des itérations successives de ces graphes) remplit l'axe des vérifiés, et moins il y a d'entiers non-vérifiés disponibles sur l'axe des vérifiés. Cela amène, selon toute probabilités, à un ensemble A(n) de plus en plus grand. Or, plus l'ensemble de départ des entiers vérifiés A(n) augmente et plus la borne inférieure pour le nombre d'itérations d'un cycle non-trivial augmente, jusqu'à prendre des valeurs extrêmement grandes.

Un cycle non-trivial doit nécessairement, car il ne peut pas tendre vers 1, aller d'entiers non-vérifiés à entiers non-vérifiés. Il semble difficile d'imaginer des cycles non-triviaux d'une longueur minimale de plusieurs centaines de millions d'entiers non-vérifiés sans être égaux à un entier vérifié, dans le contexte d'un ensemble des vérifiés de plus en plus énorme, amenant à une disparition progressive des entiers non-vérifiés. Nous avons estimé dans la **partie V point 4)**, que les intervalles proches de A(n) étaient denses en vérifiés. Un éventuel cycle non-trivial devra lui aussi passer à travers ces intervalles denses en vérifiés, sans jamais être égal à un vérifié.

Nous pouvons rajouter que ces entiers vérifiés constituant la longueur minimale d'un cycle non-trivial devraient être, non pas pris au hasard, mais parfaitement ordonnés selon les proportions propres aux caractéristiques des trajectoires de Syracuse. Cela réduit d'autant les probabilités d'existence d'un tel cycle non-trivial. De plus, le **théorème 10** nous montre que s'il existe un cycle non-trivial, une infinité de trajectoires avec des cycles non-triviaux existe, et rend cette possibilité encore plus improbable. Cela voudrait dire qu'une infinité d'entiers non-vérifiés existent, ce qui est en contradiction avec le fait, montré par diverses approches probabilistes, ainsi que par les graphes inverses, que la probabilité d'existence d'entiers ne vérifiant pas la conjecture est quasi-nulle. Le caractère de plus en plus discontinu de l'ensemble des non-vérifiés, rend fortement improbable l'hypothèse de l'existence d'un cycle non-trivial d'un minimum de plusieurs centaines de millions d'itérations avec des entiers non-vérifiés. De plus, le **théorème 11**, impose à toutes les valeurs des graphes inverses des valeurs d'un hypothétique cycle non-trivial de ne jamais être égales à un entier vérifié.

Nous obtenons donc le résultat suivant :

RESULTAT 1

La probabilité pour qu'il existe une trajectoire de Syracuse avec un cycle non-trivial est extrêmement faible.

PARTIE VII

La question des trajectoires divergentes

Dans la **partie IV** nous avons montré que si toutes les trajectoires sont majorées et qu'elles n'ont pas de cycles non-triviaux, alors elles sont vérifiées. Nous venons d'examiner l'hypothèse que les trajectoires peuvent avoir des cycles non-triviaux. Nous en avons conclu que la probabilité pour qu'une trajectoire ait un cycle non-trivial est très faible.

Nous allons examiner maintenant la deuxième partie des hypothèses laissées irrésolues dans la **partie IV**, c'est à dire le fait qu'une trajectoire puisse être divergente, c'est à dire qu'elle ne puisse pas être majorée.

Nous pouvons remarquer dans un premier temps que s'il existe une trajectoire divergente, alors il en existe une infinité. Considérons le graphe inverse de n'importe quelle valeur de la trajectoire divergente. Toutes les valeurs de ce graphe inverse auront des trajectoires qui mèneront à la trajectoire divergente. De plus, le graphe inverse de n'importe quel entier est infini (**théorème 6**).

Nous obtenons donc le résultat suivant :

THEOREME 12

S'il existe une trajectoire divergente, alors il existe une infinité de trajectoires divergentes.

S'il existe une trajectoire divergente, cela veut dire que chaque majorant atteint par la trajectoire est dépassé par d'autres majorants, et cela d'une manière infinie. Cela veut dire que la trajectoire évolue dans les limites d'un majorant, comme nous l'avons vu dans la **partie IV**, sans être égale à un vérifié, puis dépasse ce majorant, évolue dans les limites d'un nouveau majorant sans être égale à un entier vérifié, puis le dépasse de nouveau, et ainsi de suite d'une manière infinie. Le problème que nous avions à trouver plusieurs centaines de millions d'entiers non-vérifiés pour espérer réunir les conditions minimales pour l'existence d'un cycle non-trivial, est le même pour trouver une infinité de valeurs non-vérifiées en ce qui concerne une trajectoire divergente. Là encore, il y a une difficulté : réaliser une infinité d'itérations en allant d'entiers non-vérifiés à entiers non-vérifiés, sans jamais être égale à un entier vérifié, dans le contexte d'un ensemble des vérifiés en augmentation constante, et avec une probabilité quasi nulle d'existence d'entiers ne vérifiant pas la conjecture. C'est un peu comme si l'ensemble de plus en plus discontinu des entiers non-vérifiés ne pouvait pas « soutenir » une infinité de trajectoires divergentes qui nécessiteraient une infinité de non-vérifiés, sans jamais être égales à un entier vérifié.

Nous arrivons donc au résultat suivant :

RESULTAT 2

La probabilité pour qu'il existe une trajectoire de Syracuse divergente est extrêmement faible.

PARTIE VIII

Conclusion

Nous avons montré dans la **partie IV** que toutes les trajectoires de Syracuse qui étaient majorées et sans cycle non-trivial étaient vérifiées sans exception. Nous avons vu dans les **parties VI et VII** que les probabilités pour qu'une trajectoire de Syracuse soit divergente ou ait un cycle non-trivial sont extrêmement faibles. Ceci est dû principalement à deux causes : d'une part, les graphes inverses, et d'autres approches probabilistes, nous montrent que les probabilités d'existence d'entiers ne vérifiant pas la conjecture de Syracuse sont faibles. D'un autre côté, un seul entier ne vérifiant pas la conjecture implique l'existence d'une trajectoire divergente, ou d'un cycle non-trivial, et donc l'existence d'une infinité d'entiers ne vérifiant pas la conjecture. En effet tous les éléments des graphes inverses de tous les termes du cycle non-trivial ou de la trajectoire divergente devront être non-vérifiés. Ce décalage entre un nombre estimé assez faible d'entiers ne vérifiant pas la conjecture, et la nécessité d'avoir une infinité d'entiers non-vérifiés pour soutenir une hypothétique trajectoire divergente ou un hypothétique cycle non-trivial d'une longueur minimale de plusieurs centaines de millions de termes, amène à penser que les probabilités d'existence d'une trajectoire divergente ou d'un cycle trivial sont extrêmement faibles.

En résumé, nous sommes face à deux hypothèses contradictoires :

a) Première hypothèse, il n'y a pas de trajectoires divergentes, ni de cycles non-triviaux, toutes les trajectoires rentrent dans le cadre de la **partie IV**, et sont donc vérifiées. La conjecture de Collatz est vérifiée, nous avons la confirmation que G(1) = N*, car toutes les trajectoires tendent vers 1.

b) Deuxième hypothèse, il y a au moins un entier non-vérifié, ce qui veut dire qu'aucun élément des graphes inverses d'un minimum de plusieurs centaines de millions d'entiers non-vérifiés ne doit être égal à un entier vérifié.

L'étude que nous venons de mener montre que les probabilités que G(1) = N* sont extrêmement fortes, et que les probabilités de l'existence d'une suite divergente ou d'un cycle non-trivial sont extrêmement faibles.

L'intérêt théorique de l'axe des vérifiés est que de multiples recherches peuvent se greffer sur ce concept. Nous avons vu par exemple que certains résultats sur la structure des cycles non-triviaux permettent de placer des contraintes qui rendent l'existence de tels cycles difficiles. En croisant divers résultats, qu'ils viennent du domaine des graphes inverses, d'approches probabilistes, de vérifications statistiques, ou de bien d'autres domaines, nous pouvons progresser dans notre certitude que la conjecture est vérifiée. Moins il y a de cas théoriques ne vérifiant pas la conjecture, et moins les trajectoires de ces cas théoriques de plus en plus rares auront d'espace libre sur l'axe des vérifiés pour se mouvoir : en effet, il faut déjà, dans l'état actuel de nos connaissances, un très grand nombre d'entiers non-vérifiés pour pouvoir espérer mettre en place une trajectoire possiblement non-vérifiée.

L'intérêt des graphes inverses est que ceux-ci sont accessibles au calcul. Un programme tout simple permettrait de calculer les graphes inverses de n'importe quel entier. Il semble beaucoup plus intéressant, au niveau de la rapidité de calcul, de calculer directement les valeurs successives de G(1), plutôt que de vérifier une à une toutes les trajectoires de l'intervalle [1 ; 10^{20}]. En effet, moins de 3000 itérations de G(1) permettent de vérifier l'ensemble des entiers de l'intervalle [1 ; 10^{20}] (voir **partie V point 4)**).

En réalisant ce travail d'itérations sur G(1), nous pourrions examiner la vitesse d'évolution de cet ensemble, et notamment de l'ensemble de départ des entiers vérifiés A(n) qui le constitue. Nous pourrions également examiner la structure de l'ensemble des entiers non-vérifiés au dessus de la borne supérieure de A(n) afin d'évaluer les difficultés qu'auraient les trajectoires non-vérifiées restantes, à s'extraire de la zone dense en vérifiés qui se situe au dessus de A(n), sans être égales à une valeur vérifiée.

De plus en plus de démonstrations tendent à prouver que la conjecture est vérifiée. Il s'agirait de mettre bout à bout toutes ces démonstrations pour acquérir la quasi-certitude que la conjecture de Collatz est vérifiée. Cela serait une solution approchée au problème en attendant une hypothétique solution définitive, sûre et certaine. Il est cependant très possible que le problème de Collatz soit indécidable (voir J.H Conway : *on unseattleable arithmetical problems*). Cela expliquerait peut-être pourquoi aucune solution définitive à ce problème n'est trouvée depuis plus de 80 ans.

Références :

T.Tao : *Almost all orbits of the Collatz map attain almost bounded values.*

E.Rosendaal : *ericr.nl « 3x+1 delay records .»*

S. Eliahou *: The 3x+1 problem : new lower bounds on non-trivial cycle lengths.*

L-O. Pochon, A. Favre *: La suite de Syracuse, un monde de conjectures.* 2017 HAL-01593181v1.

J.H Conway : *on unsettleable arithmetical problems*. CDOI 10.4169/Amer.Math.Monthly.120.03.192.