Proof of the Collatz Theorem

Meier Franz-Xaver, Stuttgart

June 21, 2010

Abstract

Collatz theorem on $\mathbb{N}$ is an easy consequence of an equivalent problem on the monoid M generated by the five functions $\alpha, \beta, \gamma, \varphi, \psi : \mathbb{N} \to \mathbb{N}$ with $\alpha(x) = 3x + 1$, $\beta(x) = 3x + 2$, $\gamma(x) = 3x$, $\psi(x) = 2x$, $\varphi(x) = 2x+1$. The proof of Collatz theorem uses an astonishing involutive monoid automorphism $\tau : M \to M$, that maps "odd" to "even". Have fun !

1 Binaries, Ternaries, Mixed Numeral System

Notation: Given a subset A of a monoid , we denote by $\mathfrak{M}(A)$ the submonoid generated by A . $\mathfrak{M}(A)$ consists of all finite products of elements in A .

We define functions $\alpha, \beta, \gamma, \varphi, \psi : \mathbb{N} \to \mathbb{N}$:

$$\begin{aligned} \alpha(x) &:= 3x + 1 \\ \beta(x) &:= 3x + 2 \\ \gamma(x) &:= 3x \\ \varphi(x) &:= 2x + 1 \\ \psi(x) &:= 2x \end{aligned}$$

As usual the set of all functions $\mathbb{N} \to \mathbb{N}$ together with composition forms a noncommutative monoid[1]. On the submonoid $M := \mathfrak{M}(\{\alpha, \beta, \gamma, \varphi, \psi\}) \subseteq \mathbb{N}^{\mathbb{N}}$ generated by the elements $\alpha, \beta, \gamma, \varphi, \psi$ we easily verify the following six equations[2]:

$$\begin{array}{llllll} \alpha\varphi = \psi\beta & (1) & \qquad & \alpha\psi = \varphi\gamma & (2) \\ \beta\varphi = \varphi\beta & (3) & & \gamma\psi = \psi\gamma & (4) \\ \gamma\varphi = \varphi\alpha & (5) & & \beta\psi = \psi\alpha & (6) \end{array}$$

e.g. : $\alpha(\varphi(x)) = 3(2x + 1) + 1 = 6x + 4 = 2(3x + 2) = \psi(\beta(x))$.

[1] The composition $f \circ g$ is written in multiplicative notation fg . On M we only use this monoid structure, there should be no confusion with the usual multiplication on $\mathbb{N}$

[2] More precise: $f\alpha\varphi g = f\psi\beta g \ \ \forall f , g \in M$ ...

The following monoid automorphism τ obviously is an **involution** :

$$\begin{array}{rccc} \tau: & M & \xrightarrow{\sim} & M \\ & \alpha & \mapsto & \alpha \\ & \beta & \mapsto & \gamma \\ & \gamma & \mapsto & \beta \\ & \varphi & \mapsto & \psi \\ & \psi & \mapsto & \varphi \end{array}$$

It transforms the above equations (1) $\leftrightarrow$ (2) , (3) $\leftrightarrow$ (4) and (5) $\leftrightarrow$ (6) into each other. Furthermore τ induces monoid automorphisms on the submonoids

$$T := \mathfrak{M}\{\alpha, \beta, \gamma\} \subseteq M$$

(the “ternaries”) and

$$B := \mathfrak{M}\{\varphi, \psi\} \subseteq M$$

(the “binaries”).

The **length** of an element $m = m_1 m_2 \cdots m_n \in M$ with $m_i \in \{\alpha, \beta, \gamma, \varphi, \psi\}$ is defined by $lg(m) := n$. It is well-defined, because (1) - (6) are length preserving. For $k \in \mathbb{N} \cup \{0\}$ we set:

$$\begin{aligned} M_k &:= \{m \in M \ : \ lg(m) = k\} \\ B_k &:= \{m \in B \ : \ lg(m) = k\} \\ T_k &:= \{m \in T \ : \ lg(m) = k\} \end{aligned}$$

2 Equivalence relations on M

For each $a \in M$ the monoid multiplication on M induces two maps :

$$\begin{array}{rccc} \lambda_a: & M & \longrightarrow & M \\ & m & \mapsto & am \end{array}$$

$$\begin{array}{rccc} \rho_a: & M & \longrightarrow & M \\ & m & \mapsto & ma \end{array}$$

With this notation we define relations on M (i.e. subsets of $M \times M$) :

$$\begin{aligned} R_a^\lambda &:= \{(m \, , \, \lambda_a(m)) : \ m \in M\} \\ R_a^\rho &:= \{(m \, , \, \rho_a(m)) : \ m \in M\} \\ R_\tau &:= \{(m \, , \, \tau(m)) : \ m \in M\} \\ R_{\alpha\varphi} &:= \{(\lambda_\varphi(m) \, , \, \lambda_{\alpha\varphi}(m)) : \ m \in M\} \\ R_{\alpha\psi} &:= \{(\lambda_\psi(m) \, , \, \lambda_{\alpha\psi}(m)) : \ m \in M\} \end{aligned}$$

These relations generate equivalence relations on M :

$$E_a^{\lambda} := \text{equivalence relation on } M \text{ generated by } R_a^{\lambda}$$

$$E_a^{\rho} := \text{equivalence relation on } M \text{ generated by } R_a^{\rho}$$

$$E_{\tau} := \text{equivalence relation on } M \text{ generated by } R_{\tau}$$

$$E_{\alpha\varphi} := \text{equivalence relation on } M \text{ generated by } R_{\alpha\varphi}$$

$$E_{\alpha\psi} := \text{equivalence relation on } M \text{ generated by } R_{\alpha\psi}$$

Finally we set ($\vee$ is the join of equivalence relations) :

$$E_1 := E_{\alpha\varphi} \vee E_{\psi}^{\lambda} \vee E_{\psi}^{\rho} \vee E_{\gamma}^{\rho}$$

$$E_2 := E_{\alpha\psi} \vee E_{\varphi}^{\lambda} \vee E_{\varphi}^{\rho} \vee E_{\beta}^{\rho}$$

$$E_3 := E_1 \cap E_2$$

$$E_4 := E_1 \vee E_2$$

Furthermore let $[m]_1$, $[m]_2$, $[m]_3$,$[m]_4$, $[m]_{\tau}$ denote the equivalence classes of E_1 , E_2 , E_3 , E_4 , E_{τ} . We will write (for $i = 1, \ldots, 4$):

$$x \underset{i}{\sim} y \text{ instead of } (x\,,\,y) \in E_i$$

$$x \underset{\tau}{\sim} y \text{ instead of } (x\,,\,y) \in E_{\tau}$$

By definition of E_1 , E_2 , E_3 , E_4 and E_{τ} it is clear , that for all $m \in M$:

$$\begin{array}{ccccccc} \alpha\varphi m & \underset{1}{\sim} & \varphi m & \qquad & \alpha\psi m & \underset{2}{\sim} & \psi m \\ \psi m & \underset{1}{\sim} & m & & \varphi m & \underset{2}{\sim} & m \\ m & \underset{1}{\sim} & m\psi & & m & \underset{2}{\sim} & m\varphi \\ m & \underset{1}{\sim} & m\gamma & & m & \underset{2}{\sim} & m\beta \end{array}$$

$$a \underset{1}{\sim} b \Longleftrightarrow \tau\,(a) \underset{2}{\sim} \tau\,(b) \qquad \forall\, a\,,\, b \in M$$

$$a \underset{3}{\sim} b \Longleftrightarrow \tau\,(a) \underset{3}{\sim} \tau\,(b) \qquad \forall\, a\,,\, b \in M$$

It is not evident, that $\underset{1}{\sim}$ or $\underset{2}{\sim}$ are compatible with the monoid structure on M. We want to prove it in the following theorem :

Theorem 2.1 (Collatz Theorem)

(a) The quotient set $M/\underset{1}{\sim} = \{[\alpha]_1\,,\,[\gamma]_1\}$ *contains the two elements* $[\gamma]_1 = \mathfrak{M}\{\gamma,\psi\}$ *and* $[\alpha]_1 = M - \mathfrak{M}\{\gamma,\psi\}$

(b) $\underset{1}{\sim}$ *is compatible with the monoid structure and the quotient monoid* $M/\underset{1}{\sim}$ *is isomorph to the monoid* ($\mathbb{Z}/2\,,\,\cdot$) *with neutral element* $[\gamma]_1$ *and the idempotent* $[\alpha]_1$.

(c) The quotient set $M/\underset{2}{\sim} = \{[\alpha]_2\,,\,[\beta]_2\}$ *contains the two elements* $[\beta]_2 = \mathfrak{M}\{\beta,\varphi\}$ *and* $[\alpha]_2 = M - \mathfrak{M}\{\beta,\varphi\}$

(d) $\underset{2}{\sim}$ *is compatible with the monoid structure and the quotient monoid* $M/\underset{2}{\sim}$ *is isomorph to the monoid* ($\mathbb{Z}/2\,,\,\cdot$) *with neutral element* $[\beta]_2$ *and the idempotent* $[\alpha]_2$.

Proof : We define multiplicatively closed subsets of M :

$$S_1 := \mathfrak{M}\{\gamma,\psi\} \text{ and } S_2 := \mathfrak{M}\{\beta,\varphi\}$$

$$C_1 := M - S_1 \text{ and } C_2 := M - S_2 \text{ (see lemma 2.2 below)}$$

$$\Delta := C_1 \cap C_2 = (M - S_1) \cap (M - S_2) = M - S_1 - S_2$$

and prove (a) and (c) in 4 steps:

Step 1: Start with the elements of S_1 and S_2 .

i) $m \underset{1}{\sim} \gamma \;\forall\, m \in S_1$, because : $m \underset{1}{\sim} m\psi$ and $m \underset{1}{\sim} m\gamma \quad \forall\, m \in M$

ii) $m \underset{2}{\sim} \beta \;\forall\, m \in S_2$, because : $m \underset{2}{\sim} m\varphi$ and $m \underset{2}{\sim} m\beta \quad \forall\, m \in M$

iii) $\forall\, m \in S_2$ *with* $lg(m) > 0 \quad \exists\, \delta \in \Delta :\; m \underset{1}{\sim} \delta$, because:

$$\beta^\nu\varphi^\mu = \varphi\beta^\nu\varphi^{\mu-1} \underset{1}{\sim} \alpha\varphi\beta^\nu\varphi^{\mu-1} = \psi\beta\beta^\nu\varphi^{\mu-1} \underset{1}{\sim} \beta^{\nu+1}\varphi^{\mu-1} \underset{(inductive)}{\dots}$$

$$\underset{1}{\sim} \beta^{\nu+\mu} \underset{1}{\sim} \beta^{\nu+\mu}\psi = \psi\alpha^{\nu+\mu} \underset{1}{\sim} \alpha^{\nu+\mu} \in \Delta.$$

iv) $\forall\, m \in S_1$ *with* $lg(m) > 0 \quad \exists\, \delta \in \Delta :\; m \underset{2}{\sim} \delta$, because:

$$\gamma^\nu\psi^\mu = \psi\gamma^\nu\psi^{\mu-1} \underset{2}{\sim} \alpha\psi\gamma^\nu\psi^{\mu-1} = \varphi\gamma\gamma^\nu\psi^{\mu-1} \underset{2}{\sim} \gamma^{\nu+1}\psi^{\mu-1} \underset{(inductive)}{\dots}$$

$$\underset{2}{\sim} \gamma^{\nu+\mu} \underset{2}{\sim} \gamma^{\nu+\mu}\varphi = \varphi\alpha^{\nu+\mu} \underset{2}{\sim} \alpha^{\nu+\mu} \in \Delta.$$

Step 2: According to step 1 it is sufficient to prove:

$$m \underset{1}{\sim} \alpha \text{ and } m \underset{2}{\sim} \alpha \;\; \forall\, m \in \Delta = M - S_1 - S_2$$

or equivalently (because of $E_3 \mid_\Delta = E_1 \mid_\Delta \cap E_2 \mid_\Delta$):

$$m \underset{3}{\sim} \alpha \;\; \forall\, m \in \Delta \qquad (*)$$

We achieve this in two additional steps:

Step 3 : The goal of step 3 is to prove $(*)$ modulo $\underset{\tau}{\sim}$:

τ is an involutive monoid automorphism . Therefore the equivalence classes of E_τ contain at most two elements : $[m]_\tau = \{m\,,\,\tau(m)\}$. Let π denote the projection $\pi : M \to M/\underset{\tau}{\sim}$. We set $\overline{M} := M/\underset{\tau}{\sim}$,

$\overline{\Delta} := \pi\,(\Delta)$ and:

$$\overline{m} := \pi\,(m) = [m]_\tau \text{ for } m \in M$$

The image $\overline{E_3} := (\pi \mid_\Delta \times \pi \mid_\Delta)\,(E_3 \mid_\Delta)$ of the equivalence relation $E_3 \mid_\Delta$ under the map

$$\pi \mid_\Delta \times \pi \mid_\Delta : \; \Delta \times \Delta \to \overline{\Delta} \times \overline{\Delta}$$

is an equivalence relation on $\overline{\Delta}$. We write $\overline{x} \approx \overline{y}$ instead of $(\overline{x}\,,\,\overline{y}) \in \overline{E_3}$ for all $\overline{x}\,,\,\overline{y} \in \overline{\Delta}$. With this notation we want to prove :

$$\overline{m} \approx \overline{\alpha} \;\; \forall\, \overline{m} \in \overline{\Delta}$$

According to lemma 2.3 below we know, that $E_4 \mid_\Delta = (E_1 \vee E_2) \mid_\Delta = \Delta \times \Delta$ and therefore $m \underset{4}{\sim} \alpha \;\; \forall\, m \in \Delta$. By definition of the join we find for each $m \in \Delta$ a chain:

$$\alpha = m_0 \underset{1}{\sim} m_1 \underset{2}{\sim} m_2 \underset{1}{\sim} m_3 \underset{2}{\sim} m_4 \underset{1}{\sim} \ldots m_r = m$$

or

$$\alpha = m_0 \underset{2}{\sim} m_1 \underset{1}{\sim} m_2 \underset{2}{\sim} m_3 \underset{1}{\sim} m_4 \underset{2}{\sim} \ldots m_r = m$$

with $m_i \in \Delta$. In $\overline{\Delta}$ this means:

$$\overline{\alpha} = \overline{m}_0 \approx \overline{m}_1 \approx \overline{m}_2 \approx \overline{m}_3 \approx \overline{m}_4 \approx \ldots \overline{m}_r = \overline{m}$$

Step 4 : In this final step we lift the result from $\overline{\Delta}$ to Δ:

For $m \in \Delta$ we proved in step 3 , that $\overline{m} \approx \overline{\alpha}$. By definition of $\approx$ this means , that $\tau(m) \sim \tau(\alpha)$ ***or*** $m \sim \alpha$ ***or*** $\tau(m) \sim \alpha$ ***or*** $m \sim \tau(\alpha)$.

In all four cases we conclude, that $m \sim \alpha$, because $\tau(\alpha) = \alpha$ and :

$$a \sim b \Longleftrightarrow \tau(a) \sim \tau(b) \qquad \forall\, a\,,\, b \in M.$$

This completes the proof of (a) and (c) . The statements (b) and (d) are an easy consequence.

q.e.d.

Lemma 2.2 *$C_1 = M - S_1$ and $C_2 = M - S_2$ are multiplicatively closed subsets of M .*

Proof : Define a monoid homomorphism χ , well-defined because of (1) - (6) :

$$\begin{array}{rccc} \chi: & M & \stackrel{\sim}{\longrightarrow} & (\mathbb{Z}\,,\,\cdot) \\ & \alpha & \mapsto & 0 \\ & \beta & \mapsto & 0 \\ & \gamma & \mapsto & 1 \\ & \varphi & \mapsto & 0 \\ & \psi & \mapsto & 1 \end{array}$$

Clearly $S_1 = \chi^{-1}(1)$ and $C_1 = \chi^{-1}(0)$ and $C_2 = \tau(C_1)$. Therefore C_1 and C_2 are multiplicatively closed.

q.e.d.

Lemma 2.3 *a) $E_4 = E_1 \vee E_2 = M \times M$ and b) $E_4 \mid_\Delta = (E_1 \vee E_2) \mid_\Delta = \Delta \times \Delta$*

Proof : a) inductive step:
$\gamma m \underset{4}{\sim} \varphi\gamma m = \alpha\psi m \underset{4}{\sim} \psi m \underset{4}{\sim} m \underset{4}{\sim} \varphi m \underset{4}{\sim} \alpha\varphi m = \psi\beta m \underset{4}{\sim} \beta m \;\; \forall m \in M$
and finally $\alpha m \underset{4}{\sim} m \;\; \forall m \in M$, because $\alpha\psi m \underset{4}{\sim} m \underset{4}{\sim} \alpha\varphi m \; \forall m \in M$.
b) TBD

q.e.d.

We derive the classical Collatz theorem on $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ from the above Collatz theorem on M , using the surjective map:

$$\begin{array}{rccc} \vartheta : & C_1 & \rightarrow & \mathbb{N} \\ & f & \mapsto & f(0) \end{array}$$

The image $E_{col} := (\vartheta \times \vartheta)(E_1 |_{C_1})$ of the equivalence relation $E_1 |_{C_1}$ under the map:

$$\vartheta \times \vartheta \, : \, C_1 \times C_1 \rightarrow \mathbb{N} \times \mathbb{N}$$

is an equivalence relation on $\mathbb{N}$ (the ***Collatz equivalence relation***) .
We write $n \underset{col}{\sim} m$ instead of $(n\, ,\, m) \in E_{col}$ for all $n\, ,\, m \in \mathbb{N}$. By definition of $\underset{col}{\sim}$ it is clear , that $\forall\, n \in \mathbb{N}$:

$$\alpha\,(\varphi\,(n)) \underset{col}{\sim} \varphi(n)$$

$$\psi\,(n) \underset{col}{\sim} n$$

and E_{col} is the smallest equivalence relation on $\mathbb{N}$ generated by the relations:

$$\{(\alpha\,(\varphi\,(n))\, ,\, \varphi(n)) \, : \, n \in \mathbb{N}\} \subseteq \mathbb{N} \times \mathbb{N}$$
$$\{(\psi\,(n)\, ,\, n) \, : \, n \in \mathbb{N}\} \subseteq \mathbb{N} \times \mathbb{N}$$

Corollary 2.4 (Classical Collatz on $\mathbb{N}$) *All elements in* $\mathbb{N}$ *are Collatz equivalent to 1:*

$$\forall n \in \mathbb{N} : \; n \underset{col}{\sim} 1$$

Proof : The map ϑ induces a surjective map $\overline{\vartheta} : C_1/\!\sim \; \longrightarrow \mathbb{N}/\underset{col}{\sim}$, which makes the following diagram commutative:

$$\begin{array}{ccc} C_1 & \longrightarrow & \mathbb{N} \\ \downarrow & & \downarrow \\ C_1/\!\sim & \longrightarrow & \mathbb{N}/\underset{col}{\sim} \end{array}$$

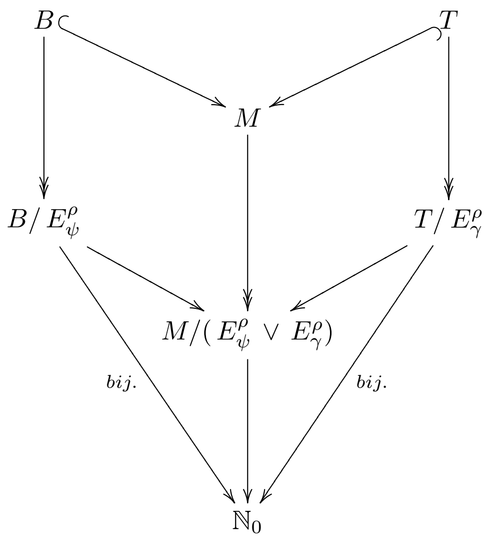
$\Longrightarrow \qquad 1 = card\,(C_1/\!\sim) \geq card\left(\mathbb{N}/\underset{col}{\sim}\right)$

$\Longrightarrow \qquad \mathbb{N}/\underset{col}{\sim}$ contains only one element, which of course is $[1]_{col}$. *q.e.d.*

3 Conversion: Binaries $\leftrightarrow$ Ternaries

Each number $n \in \mathbb{N}$ has a unique representation in the binary and the ternary numeral system, i.e. $\exists!\, f \in B - B\psi\,,\ g \in T - T\gamma$ with:

$$\begin{array}{ll} n = f\psi^{\nu}\,(0) & \forall\, \nu \geq 0 \\ n = g\gamma^{\mu}\,(0) & \forall\, \mu \geq 0 \end{array}$$



This commutative diagram (of sets) and the following proposition describe the conversion between different numeral systems.

Proposition 3.1 *(Binary-Ternary-Conversion)*

(a) $\forall\, f \in B_k\ \exists!\, g \in T_k$ *with* $f\gamma^k = g\psi^k$
(b) $\forall\, g \in T_k\ \exists!\, f \in B_{2k}$ *with* $f\gamma^k = g\psi^{2k}$
(c) $\forall\, f \in B_k\ \exists!\, g \in T_k$ *with* $f\beta^k = g\varphi^k$
(d) $\forall\, g \in T_k\ \exists!\, f \in B_{2k}$ *with* $f\beta^k = g\varphi^{2k}$

Proof : We prove here only $k = 1$ (derive the general case $k > 1$ inductive). Existence:

(a) and (c) :

$$\begin{array}{c} \varphi\gamma = \alpha\psi \\ \psi\gamma = \gamma\psi \end{array} \quad \text{and} \quad \begin{array}{c} \psi\beta = \alpha\varphi \\ \varphi\beta = \beta\varphi \end{array}$$

(b) and (d) :

$$\begin{array}{c} \alpha\psi^2 = \varphi\psi\gamma \\ \beta\psi^2 = \psi\varphi\gamma \\ \gamma\psi^2 = \psi\psi\gamma \end{array} \quad \text{and} \quad \begin{array}{c} \alpha\varphi^2 = \psi\varphi\beta \\ \gamma\varphi^2 = \varphi\psi\beta \\ \beta\varphi^2 = \varphi\varphi\beta \end{array}$$

Uniqueness: Is a consequence of the uniqueness of binary (resp. ternary) representation of numbers (use the map $M \to \mathbb{N}_0$, $f \mapsto f(0)$)

q.e.d.

TBD: Define functions:

$$\begin{array}{rccc} \alpha_{n,k} : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & x & \mapsto & nx + k \end{array}$$

and study the monoid :

$$M := \mathfrak{M}\left(\{\alpha_{n,0}\,,\,\dots\,,\,\alpha_{n,n-1}\} \cup \{\alpha_{m,0}\,,\,\dots\,,\,\alpha_{m,m-1}\}\right) \subseteq \mathbb{N}^{\mathbb{N}}$$

and the automorphism group $Aut(M)$.

MeierFranzXaver "at" gmx.de