

Pi - The Wall

Eine Monte-Carlo-Simulation zum Thema: Ist Pi normal ?

J.V.Schmidt, September 2015

1 Motivation: Wie zufällig ist Pi?

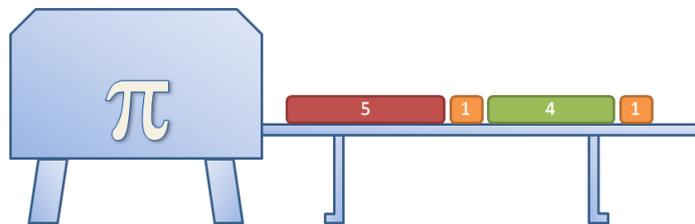
Heute soll es wieder einmal um die alte Frage gehen, ob die Zahl Pi normal ist. Anders ausgedrückt: Ist die (in unserem Falle dezimale) Ziffernfolge von Pi wirklich zufällig? Ein mathematischer Beweis dafür wurde bisher nicht vorgelegt, so dass experimentelle Untersuchungen dazu berechtigt erscheinen.

Fänden sich bei solchen Untersuchungen Abweichungen von einem erwarteten zufälligen Verhalten, wäre dies ein Indiz gegen die Zufallsvermutung.

Im weiteren präsentiere ich - spielerisch verpackt - ein derartiges Monte-Carlo-Experiment.

2 Einer hat die Absicht, eine Mauer zu errichten

Stellen wir uns eine Maschine vor, die Ziegel verschiedener Größe herstellt. Alle Ziegel sind zwar gleich hoch, doch jeder hat eine andere Breite. Die Breite der ausgeworfenen Ziegel entspricht den Ziffern der Pi-Folge. Der erste Stein ist demzufolge 1 breit, der zweite 4, der dritte 1, der vierte 5 usw.



Ein Maurer, der eine Wand errichten soll, verarbeitet die Ziegel genau in der Herstellungsreihenfolge. Er setzt zuerst die Steine der untersten Schicht aneinander.



Irgendwann wechselt unser Maurer dann zur darüberliegenden Schicht und baut hier erneut die Steine aneinander. Die unterste Schicht darf nun nicht mehr verändert, also auch kein Stein mehr angefügt werden. Natürlich möchte er eine schöne Wand bauen, bei der kein Stein übersteht.

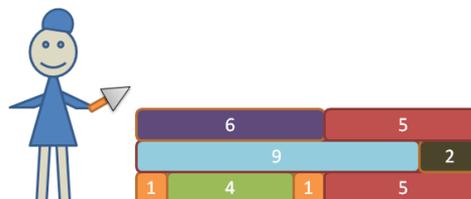
Kann ihm das gelingen?

Wie hoch schafft er es, die Mauer zu ziehen?

Und wie viele Steine liegen dabei in der untersten Schicht?

Nach kurzem Probieren findet der Maurer heraus, dass mit vier Steinen in der ersten Schicht eine Mauer der Höhe 3 möglich ist:

Hier ergeben die vier Steine der untersten Schicht in Summe 11 ($1+4+1+5$), was dann jeweils durch zwei Steine für Schicht 2 ($9+2$) und Schicht 3 ($6+5$) exakt getroffen wird. Die Folgeschicht wird dann leider mit 16 ($3+5+8$) zu lang.



Wie aber kann er nun eine noch höhere Mauer zusammensetzen?

3 Der Rechenknecht startet

Wir sollten dem Maurer nicht gram sein, wenn er nach vielen erfolglosen Versuchen die Kelle sinken lässt, denn Mauern mit Höhen größer als 3 sind schwerlich durch Probieren zu finden.

Tatsächlich zeigt die Computersuche, dass erstmals eine Mauer der Höhe 4 zu schaffen ist, wenn in der untersten Schicht 335 Steine liegen. Diese ergeben eine Mauerbreite von 1501, welche in den folgenden Schichten jeweils exakt durch 338, 327 und 321 Steine eingestellt wird.

Offensichtlich ist es nicht einfach, bei zufällig gelieferten Steinen die Mauerbreite einzuhalten. Um uns ein Bild davon zu verschaffen, lassen wir einen Computer die Berechnungen durchführen. Das Programm liest Zahl für Zahl die Pi-Folge aus und setzt die Mauerschichten zusammen.

Die folgende Tabelle zeigt, wann jeweils erstmals eine Mauer einer bestimmten Höhe erreicht wird.

Höhe der Mauer	Breite der Mauer	Anzahl Steine in der 1. Mauerlinie
1	1	1
3	11	4
2	41	10
4	1.492	335
5	1.501	336
6	8.418	1.856
7	78.290	17.414
8	845.067	188.006
11	2.180.542	484.532
9	3.317.760	737.283
10	12.890.216	2.863.389
14	19.529.322	4.338.015

Tatsächlich: um eine Mauer der Höhe $h > 5$ zu errichten, braucht man richtig viele Steine. Warum ist das so?

4 Eine Modell finden, das weiterhilft

Um besser zu verstehen, was hier vor sich geht, benötigen wir ein mathematisches Modell der Situation. Dieses sollte im Idealfall aussagen, mit welcher Wahrscheinlichkeit eine Folgeschicht der Mauer die gleiche Breite hat wie die darunterliegende Schicht.

Wie können wir vorgehen?

Zunächst können wir feststellen, dass die unterste Schicht bei höheren Mauern sehr schnell richtig breit wird und viele Steine enthält. Bezeichnen wir die Anzahl der Steine, aus denen die Mauerschicht i besteht, mit s_i , so werden wir im Folgenden den Fall

$$s_i \gg 1 \quad (1)$$

betrachten.

Anders ausgedrückt: es geht zunächst erst einmal um sehr breite Mauern. Auf 'schmale' Mauern komme ich am Ende zu sprechen.

Jeder Stein k hat seine individuelle Breite b_k . Die Summe aller Steinbreiten einer Schicht i ergibt die Mauerbreite B_i der Schicht:

$$B_i = \sum_{b_k \in i} b_k \quad (2)$$

Damit eine Mauer der Höhe h entsteht, muss also

$$B_1 = B_2 = B_3 = \dots = B_h \quad (3)$$

gelten.

In unserer Betrachtung können wir die Ziffer 0 in der Zahlenfolge und damit alle Steine der Breite 0 ignorieren, denn diese verbreitern ja die Schicht nicht.

Entscheidend dafür, ob eine Schicht die Sollbreite B_1 der untersten Schicht erreicht, ist der letzte Stein, welcher angelegt wird. Dieser hat entweder die richtige Breite, so dass die Gesamtbreite B_1 genau getroffen wird, oder er ist zu groß und steht über. Zu klein kann dieser Stein nie sein, denn sonst wäre er ja nicht der letzte Stein.

Nun wollen wir den Begriff ZIELBEREICH für die Mauerbreite einführen.

Der Zielbereich besteht aus den Zellen Z_1, Z_2, \dots, Z_9 , die sich jeweils im Abstand 1, 2, ..., 9 vom gewünschten Schichtende B_1 befinden:



Von jedem Element Z_i aus kann ich also durch Anlegen eines einzigen Steines der Breite i genau die richtige Schichtbreite B_1 erreichen.

Da es nur Steine der Breiten 1 bis 9 gibt, muss jede Mauer bei ihrem Wachstum irgendwann im ZIELBEREICH ankommen, denn nur ein Stein mit einer Breite >9 könnte den kompletten Zielbereich überspringen.

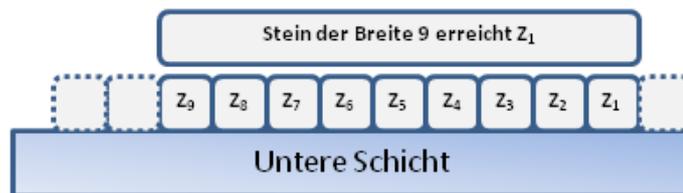
Dies wird uns als Ansatzpunkt zur Berechnung der Wahrscheinlichkeit gleicher Schichtbreiten dienen.

5 Das Modell durchrechnen: BESETZUNGSKONSTANZ

Aufgrund von Bedingung (1) endet eine Mauer, bevor sie den Zielbereich erreicht, statistisch betrachtet mit der gleichen Wahrscheinlichkeit auf jeder beliebigen Position links vom Zielbereich. Diese Eigenschaft möchte ich BESETZUNGSKONSTANZ nennen.

Dies bedeutet aber auch gleichzeitig, dass die Zellen Z_i unterschiedlich oft getroffen werden. Warum ist das so?

Schauen wir auf Zelle Z_1 : diese kann von außerhalb des Zielbereichs nur mit einem großen 9er-Stein getroffen werden.



Z_2 hingegen kann mit einem 9er- oder 8er-Stein erreicht werden.



So geht es weiter bis Zelle Z_9 , für die alle Steinbreiten von 1 bis 9 Treffer liefern können.

Insgesamt gibt es also 45 ($=1+2+\dots+8+9$) mögliche Einschlagsvarianten im Zielbereich, wobei Zelle Z_i im statistischen Mittel mit der Wahrscheinlichkeit

$$w_{i_0} = i/45 \quad (4)$$

getroffen wird.

Unsere Aufgabe besteht nun darin, für jede Zelle $Z_{i=1..9}$ des Zielbereichs zu berechnen, mit welcher Wahrscheinlichkeit w_i von dort aus die richtige Schichtbreite erreicht werden kann. Die Summe dieser (jeweils mit w_{i_0}) gewichteten Wahrscheinlichkeiten ergibt dann die gesuchte Größe zur Bewertung eines 'Mauererfolges', d.h. die Schicht endet mit der passenden Breite.

Behalten wir im Auge, dass wir eine Zufallsmaschine für die Steinherstellung verwenden: Steine mit den Breiten 1 bis 9 werden also gleich oft ausgeliefert, nämlich mit der Wahrscheinlichkeit

$$w_0 = 1/9 \quad (5)$$

Beginnen wir die Berechnungen mit der Zelle Z_1 .

Von Z_1 aus kann ich nur mit einem Stein der Breite 1 zum Ziel gelangen, denn alle anderen Steine (Breiten 2 bis 9) würden überstehen. Somit gilt

$$\begin{aligned} w_1 &= w_0 \\ &= 1/9 \end{aligned} \quad (6)$$

Betrachten wir nun Zelle Z_2 , die den Abstand 2 zum Schichtende hat.

Ein Stein der Breite 2 passt perfekt, aber auch ein Stein der Breite 1 mit einem anschließenden 'Ziellauf' ab Zelle Z_1 ist ok:

$$\begin{aligned} w_2 &= w_0 + 1/9 \cdot w_1 \\ &= 1/9 \cdot (1 + w_1) \end{aligned} \quad (7)$$

Bei Stein Z_3 sind folgende Kombinationen möglich:

- direkter Sprung ans Ende (Stein der Breite 3),
- Sprung in Zelle Z_1 (Stein der Breite 2) und dann von Z_1 ans Ende,
- Sprung in Zelle Z_2 (Stein der Breite 1) und dann von Z_2 ans Ende.

$$\begin{aligned} w_3 &= w_0 + 1/9 \cdot w_1 + 1/9 \cdot w_2 \\ &= 1/9 \cdot (1 + w_1 + w_2) \end{aligned} \quad (8)$$

Eine analoge Untersuchung der weiteren Zellen Z_4 bis Z_9 führt uns zu einer allgemeinen Rekursionsformel für die $w_{2..9}$:

$$\begin{aligned} w_i &= w_0 + 1/9 \sum_1^{i-1} w_i \\ &= 1/9 \cdot (1 + \sum_1^{i-1} w_i) \quad i \in [2..9] \end{aligned} \quad (9)$$

Entsprechend diesen Regeln lassen sich die w_i leicht berechnen, mit ihren Wichtungsfaktoren multiplizieren und addieren. So ergibt sich der gesuchte Wahrscheinlichkeitswert p dafür, dass eine wachsende Mauerschicht eine vorgegebene Breite $B \gg 1$ exakt trifft.

Die Tabelle zeigt den Berechnungsweg.

i	w_{0i} (Wichtung=i/45)	w_i	Produkt ($w_{0i} \times w_i$)
1	0,022222	0,111111	0,002469
2	0,044444	0,123457	0,005487
3	0,066667	0,137174	0,009145
4	0,088889	0,152416	0,013548
5	0,111111	0,169351	0,018817
6	0,133333	0,188168	0,025089
7	0,155556	0,209075	0,032523
8	0,177778	0,232306	0,041299
9	0,200000	0,258117	0,051623
Summe			p = 0,20

Es gilt demzufolge

$$p = 0.2 = 1/5 \quad (10)$$

Der Reziprokwert von p gibt an, wie viele Versuche wir durchschnittlich brauchen, um auf einer gegebenen Schicht eine exakt passende aufzusetzen. Im Durchschnitt überlebt also nur eine von fünf Mauern den Versuch, eine weitere Schicht aufzusetzen.

p ist gleichzeitig ein Wert für das Verhältnis der Mauern mit verschiedenen Höhen ($h > 1$):

$$p = \text{Anzahl Mauern (Höhe=h)} / \text{Anzahl Mauern (Höhe=h-1)} \quad (11)$$

$$= \text{Anzahl Mauern (Höhe} \geq h) / \text{Anzahl Mauern (Höhe} \geq h-1) \quad (12)$$

Wir haben es hier mit einem exponentiellen Abfall bei steigender Mauerhöhe zu tun, was das seltene Auftreten wirklich hoher Mauern erklärt. Der Erwartungswert für eine Mauer der Höhe h beträgt demnach

$$p_h = p^{h-1} \cdot (1 - p) \quad (13)$$

Die folgende Tabelle gibt einen Überblick zu den Erwartungswerten nach Mauerhöhen.

h Höhe der Mauer	p_h Mauern, die mindestens Höhe h erreichen	$1/p_h$ Anzahl Versuche, damit Ereignis eintritt	p_h/p_{h-1}	p^*_h Mauern, die genau Höhe h erreichen	p^*_h/p^*_{h-1}
1	100,00%	1	-	80,00%	-
2	20,00%	5	0,200	16,00%	0,200
3	4,00%	25	0,200	3,20%	0,200
4	0,80%	125	0,200	0,64%	0,200
5	0,16%	625	0,200	0,128%	0,200
6	0,032%	3.125	0,200	0,0256%	0,200
7	0,0064%	15.625	0,200	0,00512%	0,200
8	0,00128%	78.125	0,200	0,001024%	0,200
9	0,000256%	390.625	0,200	0,0002048%	0,200
10	0,0000512%	1.953.125	0,200	0,00004096%	0,200
11	0,00001024%	9.765.625	0,200	0,000008192%	0,200
12	0,000002048%	48.828.125	0,200	0,0000016384%	0,200
13	0,0000004096%	244.140.625	0,200	0,00000032768%	0,200
14	0,00000008192%	1.220.703.125	0,200	0,000000065536%	0,200
15	0,00000001638%	6.103.515.625	0,200	0,0000000131072%	0,200

6 Modell vs Messwerte

So weit, so gut.

Wir wissen nun, welche Verteilung der Mauerhöhen wir zu erwarten haben, wenn die Ziegelsteine mit zufälligen Breiten angeliefert werden und die Mauerschichten viele Steine enthalten. Es ist an der Zeit, unsere Pi-Zahlenfolge ins Rennen zu schicken.

Natürlich müssen wir eine große Anzahl von Mauern prüfen, damit das ermittelte Ergebnis statistisch signifikant ist. Man spricht in einem solchen Falle von Monte-Carlo-Simulationen: eine Situation wird immer und immer wieder mit veränderten, zufällig generierten Eingangsparametern durchgespielt; die Ergebnisse werden gesammelt. Je mehr Einzelsituationen man durchrechnet, um so zuverlässiger wird sich der Mittelwert des Ergebnisses, welcher ja aus den zufällig verteilten Einzelwerten entsteht, dem Erwartungswert der untersuchten Konstellation annähern.

Wird die Zahl Pi den Zufallstest des Mauerbaus überstehen?

Können wir den Werte $p = 0.2$ experimentell bestätigen?

6.1 Experiment I: Die wachsende Wand

Im ersten Experiment beginnen wir mit einem Stein in der untersten Reihe und versuchen, eine Mauer zu bauen - bis wir scheitern, d.h. die Folgeschicht schließt nicht mehr mit der darunterliegenden Schicht exakt ab.

Dann legen wir einen Stein mehr in die unterste Schicht und versuchen erneut, die Mauer nach oben zu ziehen. Dies wiederholen wir so oft, bis wir ausreichend Daten gesammelt haben, um den statistischen Fehler unter die gewünschte Grenze zu drücken.

Dabei merken wir uns für jeden Einzelversuch, bei welcher Mauerhöhe er endete - so ergibt sich letztlich eine Häufigkeitsverteilung der verschiedenen Mauerhöhen.

Die unterste Schicht startet stets bei der 1. Nachkommastelle von Pi, einer 1.

Das Experiment bis zu einer Steinanzahl von 1 Milliarde in der untersten Mauerschicht bringt das folgende Ergebnis:

Höhe der Mauer h	Erwartungswert Anzahl Mauern	Experimentell Anzahl Mauern A_h	p-Wert A_h/A_{h-1}
1	800.000.000	800.011.960	-
2	160.000.000	159.992.203	0,1999873
3	32.000.000	31.996.494	0,1999878
4	6.400.000	6.399.780	0,2000150
5	1.280.000	1.278.349	0,1997489
6	256.000	256.623	0,2007456
7	51.200	51.529	0,2007965
8	10.240	10.359	0,2010324
9	2.048	2.193	0,2117000
10	410	403	0,1837665
11	82	89	0,2208437
12	16	15	0,1685393
13	3	2	-
14	1	1	-
Summe		1.000.000.000	

Volltreffer!!!

Wir erhalten eine verblüffende Übereinstimmung mit der erwarteten Verteilung. Fassen wir die seltenen Ereignisse für $h \geq 12$ in einer Klasse zusammen, ergibt die Chi-Quadrat-Verteilung den Wert $\chi^2 = 19,378 < \chi_{11,0.05}^2 = 19,675$.

Mit einer Irrtumswahrscheinlichkeit $< 0,05$ entspricht das Experiment also unserer Theorie.

Ich habe dieses Experiment für verschiedene Startpunkte der Pi-Folge wiederholt, die obige Übereinstimmung mit dem Modell wurde dabei jedes Mal bestätigt.

6.2 Experiment II: Die wandernde Wand

Für den folgenden Versuch wird das vorherige Experiment leicht abgewandelt.

Wir starten nun sofort mit einer breiten unteren Schicht aus $N \gg 1$ Steinen und versuchen wieder, die Mauer hochzuziehen.

Wenn wir scheitern, wird die unterste Schicht erneut in der gleichen Breite N zusammengestellt, der Startstein wird dabei aber um eine Position nach rechts verschoben. Man könnte auch sagen: der linke Stein entfällt, am rechten Ende kommt ein Stein hinzu.

Das ist so, als würden wir bei jedem neuen Versuch den Einlesestart der Pi-Folge um eine Stelle nach rechts verschieben - die Mauer 'wandert'.

Hier das Ergebnis, wenn eine Mauer mit 10.000 Steinen in der untersten Reihe durch 100.000.000 Startpositionen der Pi-Folge wandert:

Höhe der Mauer	Erwartungswert	Experimentell	p-Wert
h	Anzahl Mauern	Anzahl Mauern	A_h/A_{h-1}
1	80.000.000	80.006.847	-
2	16.000.000	15.997.470	0,1999513
3	3.200.000	3.197.076	0,1998489
4	640.000	638.673	0,1997679
5	128.000	127.804	0,2001087
6	25.600	25.607	0,2003615
7	5.120	5.267	0,2056859
8	1.024	1.000	0,1898614
9	205	208	0,2080000
10	41	32	0,1538462
11	8	15	0,4687500
12	2	1	0,0666667
Summe		100.000.000	

Natürlich haben wir bei diesem modifizierten Experiment das gleiche Ergebnis wie im vorherigen Experiment erwartet: wir werden nicht enttäuscht.

Die Berechnung der Chi-Quadrat-Verteilung liefert den Wert $\chi^2 = 16,744 < \chi_{10,0.05} = 18,307$, also alles im grünen Bereich.

Danke, liebes Pi !

6.3 Experiment III: Schmale Wände

Bei der Herleitung unseres theoretischen Modells benötigten wir die Bedingung $s_i \gg 1$, d.h. die unterste Mauerschicht muss aus vielen Einzelsteinen bestehen.

Nur damit kann die statistische BESETZUNGSKONSTANZ VOR DEM ZIELBEREICH gewährleistet werden - eine Grundannahme unseres Modells.

Bestehen die Wände jedoch nur aus wenigen Steinen, gilt unser Modell nicht mehr.

Zur korrekten Berechnung müsste jeder Einzelfall der Mauerbreite für Schicht 1 einzeln analysiert werden. Dabei wäre jeweils zu ermitteln, auf wie viele Arten man die Gesamtbreite aus Einzelsteinen zusammensetzen kann. Dies ist mühsam, deshalb wollen wir diesen Weg nicht gehen. Was wir allerdings tun können: ein Gefühl dafür zu entwickeln, unter welchen Bedingungen unsere Hypothese der Besetzungskonstanz gilt.

Dazu lassen wir das Experiment II (Die wandernde Wand) für unterschiedliche Steinzahlen in Schicht 1 laufen. Wir starten dabei mit der Steinzahl 1 und erhöhen diese bei jedem Folgeexperiment um 1. Jedes Mal lassen wird die Mauer durch 100.000.000 Positionen wandern.

Dabei sollte sich beobachten lassen, wie sich p langsam dem Wert 0.2 nähert.

Schauen wir uns die Ergebnisse an:

Mauerbreite b (Schicht 1)	Anzahl Mauern für h=1	Anzahl Mauern für h=2	p
1	83.186.449	13.732.553	0,165082
2	80.739.540	15.425.052	0,191047
3	80.129.118	15.874.279	0,198109
4	80.016.560	15.973.570	0,199628
5	79.997.391	15.997.572	0,199976
6	79.997.820	16.001.707	0,200027
7	80.002.039	15.998.142	0,199972
8	80.002.010	15.997.903	0,199969
9	79.994.431	16.003.351	0,200056
10	80.003.091	15.999.290	0,199983
20	79.999.718	15.999.786	0,199998
50	80.003.517	15.997.263	0,199957
100	80.011.604	15.990.916	0,199857

Den p -Wert habe ich hierbei jeweils aus den ersten beiden Mauerzahlen (Mauern mit Höhe 1 bzw. 2) gebildet, da deren Erwartungswert am höchsten und somit der vermutliche Fehler am geringsten ist.

Wie man sieht, entspricht das BESETZUNGSKONSTANZ-MODELL bereits ab Steinzahlen von $s = 5$ sehr gut der Realität. Das ist mehr als überraschend!

Wer Interesse und Lust hat, kann sich gerne einmal mit der Herleitung des theoretischen p -Wertes für die kleinen Steinzahlen $s = 1$ bzw. $s = 2$ beschäftigen.

7 Resümee: Alles ziemlich normal

Wir haben uns Mühe gegeben und ein Modell entwickelt, um das Verhalten des Mauerbaus vorherzusagen. Wir haben Algorithmen erdacht, um effizient lange Zahlenketten zu lesen und zu verarbeiten. Es wurde probiert, geändert und getestet, bis schließlich alles lief. Wir mussten unseren Kopf anstrengen und hatten Spaß dabei.

Und nun sieht es wieder einmal ganz danach aus, als wäre Pi ziemlich normal.

Pi-Freund 31: Sag mal, findest Du das enttäuschend?

Pi-Freund 41: Was meinst Du?

Pi-Freund 31: Na, dass Pi so normal ist.

Pi-Freund 41: Jeder ist gern normal, nämlich so wie die andern, weil man dazugehören möchte.

Pi-Freund 31: Aber Pi ist doch nicht wie die anderen, oder?

Pi-Freund 41: Nein, natürlich nicht. Pi ist ETWAS BESONDERES.

Pi-Freund 31: Aber wieso ist Pi dann normal, wenn es nicht wie die anderen ist?

Pi-Freund 41: Weil Pi auf eine besondere Weise normal ist, eben nicht so, wie andere normal sind!

(Pi-Freund 31 denkt kurz nach.)

Pi-Freund 31: Unser Gespräch kommt mir jetzt aber irgendwie reichlich transzendent vor.

Pi-Freund 41: Mag sein.

Pi-Freund 31: In solchen Situationen hilft nur eines ...

Pi-Freund 41: Nämlich?

Pi-Freund 31: Man muss sich wieder erden.

Pi-Freund 41: Ok, einverstanden. Wollen wir?

Pi-Freund 31: Yep. Du zuerst?

Pi-Freund 41: Nein, Du!

Pi-Freund 31: Eins !

Pi-Freund 41: Vier !

Pi-Freund 31: Eins !

Pi-Freund 41: Fünf !

Pi-Freund 31: Neun !

Pi-Freund 41: Zwoooo !

Pi-Freund 31: Sechs !

Pi-Freund 41: Fünf !

.....

Und wenn sie nicht gestorben sind, dann erden sie sich heute noch.

In diesem Sinne:

Keep on piing,

Euer JVS