Elementare Erfahrungen mit chaotischem Verhalten beim NEWTON-Verfahren

Das NEWTON'sche Näherungsverfahren zur näherungsweisen Bestimmung der Nullstelle einer stetig differenzierbaren Funktion *f* beruht auf der einfachen Idee, die gute Näherung von Tangenten an Graphen zu nutzen und statt der Nullstelle des Graphen die Nullstelle einer Tangente als Näherungswert zu bestimmen.

Das Verfahren läuft in folgenden Schritten ab:

(1) Wahl eines Startpunktes $(x_0 \mid f(x_0))$ auf dem Graphen,

(2) Ermittlung der Tangente durch diesen Punkt,

(3) Bestimmung der Nullstelle x_1 dieser Tangente,

(4) Berechnung des Funktionswerts $f(x_1)$,

(5) Wahl des Punktes $(x_1 \mid f(x_1))$ als nächsten Startpunkt der Iteration.

Aus der Gleichung der Tangente *t* im Punkt $(x_0 \mid f(x_0))$ mit

$$t(x) = f'(x_0) \cdot (x - x_0) + f(x_0)$$

ergibt sich für die Nullstelle der Tangente

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

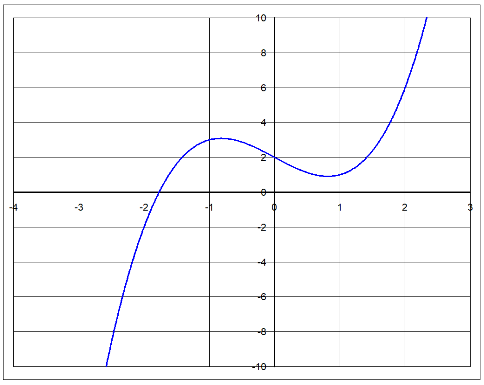
Die Wiederholung des Verfahrens erfolgt also für $n \geq 0$ mithilfe der Rekursionsformel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Die Konvergenz des Verfahrens ist nur lokal gewährleistet, d. h., der Startwert muss genügend nahe bei der gesuchten Nullstelle liegen, damit es tatsächlich konvergiert. Wenn dies der Fall ist, verdoppelt sich schließlich bei jedem Schritt die Anzahl der gültigen Stellen (sog. quadratische Konvergenzordnung).

Untersuchung einer ganzrationalen Funktion 3. Grades mit genau einer Nullstelle

In den Wikipedia-Artikeln zum NEWTON-Verfahren (deutsche und englische Version) wird im Zusammenhang mit dem Konvergenzverhalten als Beispiel die Funktion 3. Grades mit $f(x) = x^3 - 2x + 2$ betrachtet.



Entsprechende Beobachtungen können auch bei anderen Funktionen 3. Grades gemacht werden, deren Graph zwei Extrempunkte hat, welche beide oberhalb oder beide unterhalb der *x*-Achse liegen.

Solche Funktionen besitzen nur eine reelle Nullstelle, die exakt nur mithilfe der CARDANO'SCHEN Formel bestimmt werden kann.

Üblicherweise legt man vor Anwendung des NEWTON-Verfahrens eine Wertetabelle an, um das kleinste Intervall mit ganzzahligen Intervallgrenzen zu bestimmen, in dem der Vorzeichenwechsel stattfindet:

x	-3	-2	-1	0
f(*x*)	-19	-2	3	2

Mit dem Startwert $x_0 = -2$ erhält man dann mithilfe der Rekursionsvorschrift

$$x_{n+1} = x_n - \frac{x_n^3 - 2x_n + 2}{3x_n^2 - 2} = \frac{3x_n^3 - 2x_n - x_n^3 + 2x_n - 2}{3x_n^2 - 2} = \frac{2x_n^3 - 2}{3x_n^2 - 2} =: g(x_n)$$

die nächsten Folgenglieder. Deren Werte stabilisieren sich hier bereits nach vier Iterationen:

n	1	2	3	4	5
x_n	–1,8	–1,76994818...	–1,76929266...	–1,769292354...	–1,769292354...

Die Iterationen kann man mithilfe eines sog. **Spinnweb-Diagramms** veranschaulichen.

Ein solches Diagramm enthält den Graphen der Iterationsfunktion *g*, im betrachteten Beispiel also den Graphen der Funktion mit $g(x) = \frac{2x^3 - 2}{3x^2 - 2}$, und den Graphen zu $y = x$.

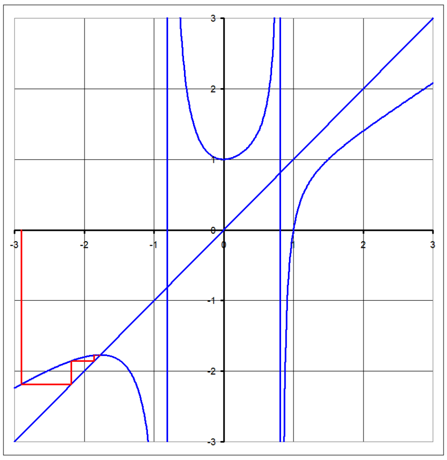
Zusätzlich zu diesen beiden Graphen zeichnet man dann wie folgt einen Streckenzug, dessen Eckpunkte abwechselnd auf den beiden Graphen liegen:

Das Einsetzen eines *x*-Werts in den Funktionsterm von *g* wird durch vertikal verlaufende Strecken veranschaulicht; durch die horizontal verlaufenden Strecken wird der Übergang zum nächsten Schritt vorgenommen: Der Funktionswert *g*(*x*) wird neuer *x*-Wert.

$(x_0 \mid g(x_0) = x_1) \rightarrow (x_1 \mid x_1) \rightarrow (x_1 \mid g(x_1) = x_2) \rightarrow (x_2 \mid x_2) \rightarrow (x_2 \mid g(x_2) = x_3) \rightarrow \ldots$

In der folgenden Abbildung ist ein solches Spinnweb-Diagramm für den Startwert $x_0 = -2{,}9$ dargestellt.

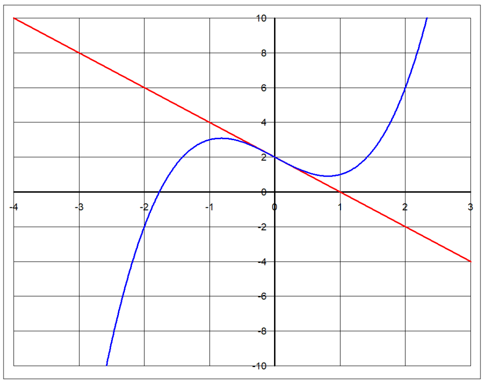
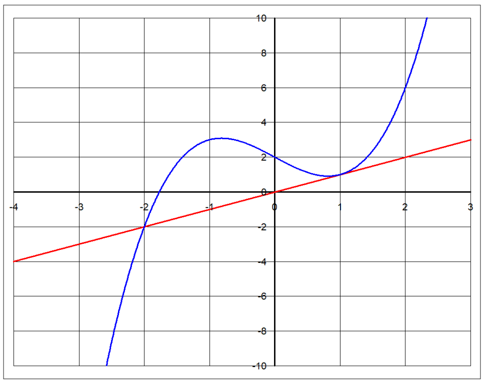
n	0	1	2	3	4	5	6
x_n	–2,9	–2,1858803...	–1,8556985...	–1,7742040...	–1,7693095...	–1,7692923...	–1,7692923...



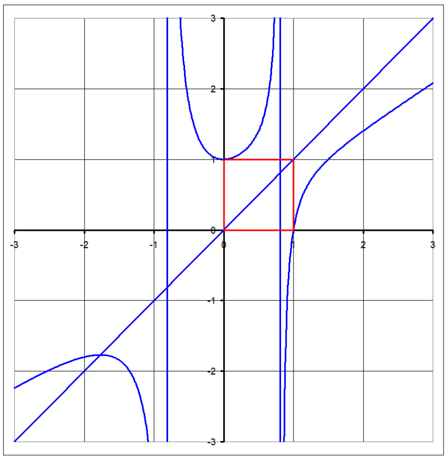
Untersuchung des Konvergenzverhaltens für verschiedene Startwerte

Eine besondere Eigenschaft der hier betrachteten Funktion *f* mit dem Symmetriezentrum $(0 \mid 2)$ ist die Eigenschaft, dass die Tangente an den Graphen an der Stelle $x = 0$ die x-Achse an der Stelle $x = 1$ schneidet und die Tangente an der Stelle $x = 1$ die x-Achse an der Stelle $x = 0$.

Daher führen die Startwerte $x_0 = 0$ und $x_0 = 1$ jeweils zu einer periodischen Folge, bei der alternierend nur die Werte 1 und 0 auftreten. Diese Folgen mit den beiden Häufungswerten 0 und 1 sind also divergent.

Das zugehörige Spinnweb-Diagramm ist der geschlossene Streckenzug, der ein Quadrat mit den Eckpunkten $(0\,|\,0)$, $(0\,|\,1)$, $(1\,|\,1)$, $(1\,|\,0)$ bildet.



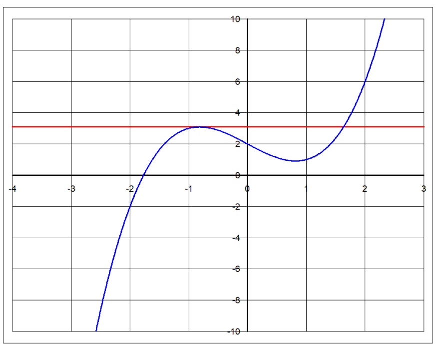
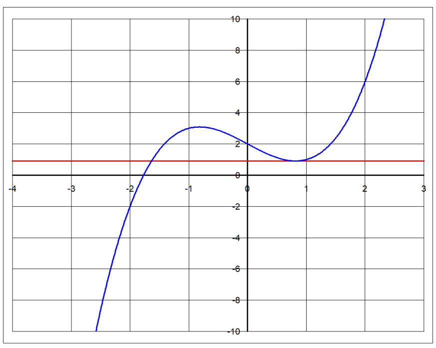
Startwerte links vom lokalen Maximum der Iterationsfunktion *g* bei $x = -1{,}7692923...$ führen in wenigen Schritten zu brauchbaren Näherungswerten für die Nullstelle der Funktion *f*.

Dass für *solche* Startwerte die Folgen konvergent sind, liegt daran, dass die Steigung der Funktion *g* hier stets kleiner ist als 1, was gemäß BANACH'schem Fixpunktsatz zu einer Kontraktion führt, d. h., dass sich der Streckenzug auf einen Punkt zusammenzieht.

Als Startwerte für eine Iteration kommen natürlich nicht die beiden lokalen Extrema der Ausgangsfunktion *f* in Frage, da die Tangente dort parallel zur *x*-Achse verläuft und keine Nullstelle besitzt.

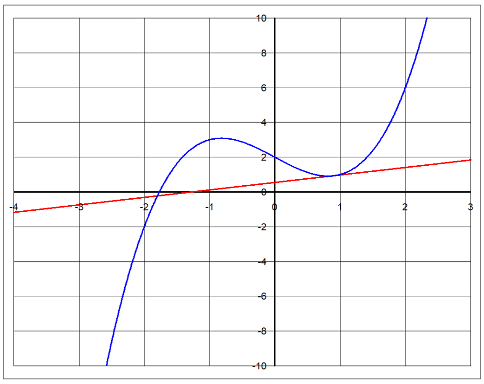
Aus der notwendigen Bedingung $f'(x) = 3x^2 - 2 = 0$ ergibt sich ein lokales Minimum von *f* bei $x = +\sqrt{\frac{2}{3}} = +0{,}81649658...$ und ein lokales Maximum von *f* an der Stelle $x = -\sqrt{\frac{2}{3}} = -0{,}81649658...$.

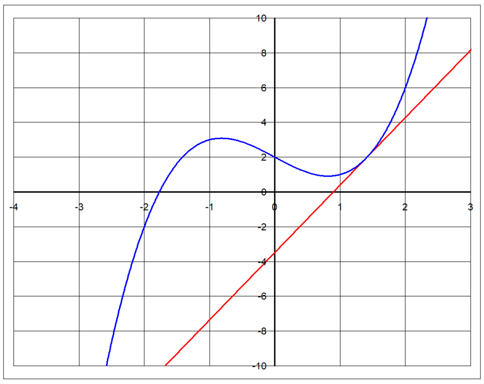
Da der Term für $f'(x)$ im Nenner der Rekursionsvorschrift auftritt, sind die Extremstellen von *f* die Polstellen von *g*.

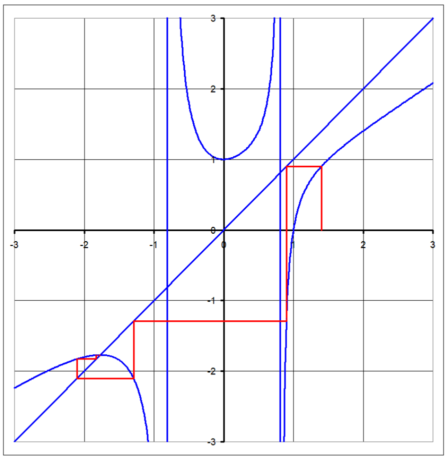
Für alle anderen Startwerte rechts vom lokalen Maximum von *g* ist allerdings das Konvergenzverhalten des Verfahrens nicht pauschal vorhersehbar.

Beispielsweise liegt die Nullstelle der Tangente im Punkt $(0{,}9 \mid f(0{,}9))$ links vom lokalen Maximum; daher ist $x_0 = 0{,}9$ als Startwert sicherlich geeignet. Und es gibt sicherlich einem geeigneten Startwert *in der Nähe* von $x_0 = 1{,}4$, von dem aus man im nächsten Schritt den Wert $x_1 = 0{,}9$ erreicht; daher ist auch ein solcher Startwert für eine konvergente Folge geeignet.





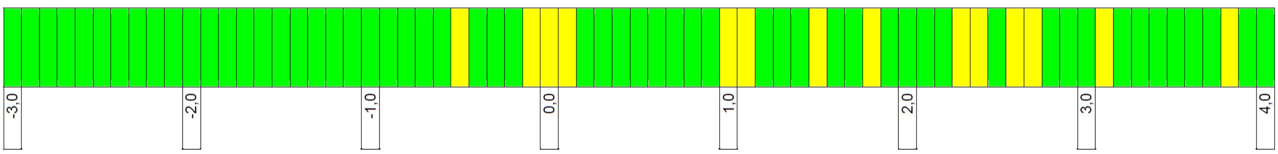
Das zum Startwert $x_0 = 1{,}4$ gehörende Spinnweb-Diagramm ist in der folgenden Abbildung zu sehen.



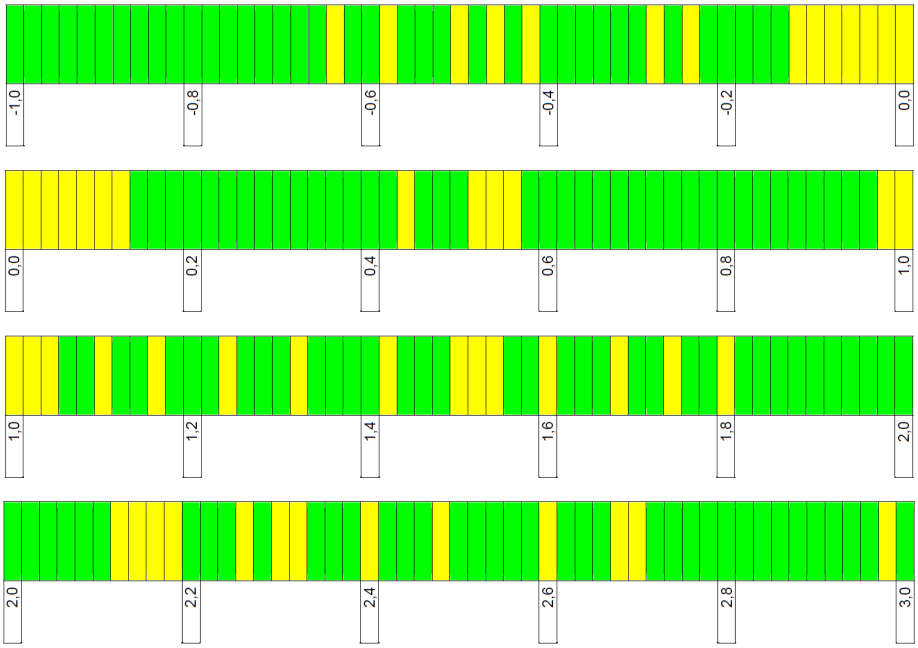
n	0	1	2	3	4	5	6	7
x_n	1,4	0,89897	-1,28878	-2,10577	-1,82920	-1,77172	-1,76930	-1,76929

Im Folgenden soll untersucht werden, bei welchen Startwerten das Verfahren zu einer konvergenten bzw. zu einer divergenten Folge führt. Die Untersuchungen hierzu wurden mithilfe der Tabellenkalkulation EXCEL (15-stellige Rechengenauigkeit) durchgeführt, wobei jeweils die ersten 100 Folgenglieder berechnet wurden. Als „konvergent" wurde eine Folge angesehen, wenn die Folgenglieder schließlich numerisch mit dem o. a. Grenzwert $-1{,}7692923...$ übereinstimmten. Bei den als „divergent" angesehenen Folgen traten schließlich abwechselnd nur die Werte 0 und 1 auf (vgl. oben).

In einem ersten Überblick wurde die Konvergenz/Divergenz mit Startwerten zwischen $x_0 = -3{,}0$ und $x_0 = +4{,}0$ untersucht – jeweils im Abstand $\Delta x = 0{,}1$; die Farben bedeuten: grün = Konvergenz, gelb = Divergenz.



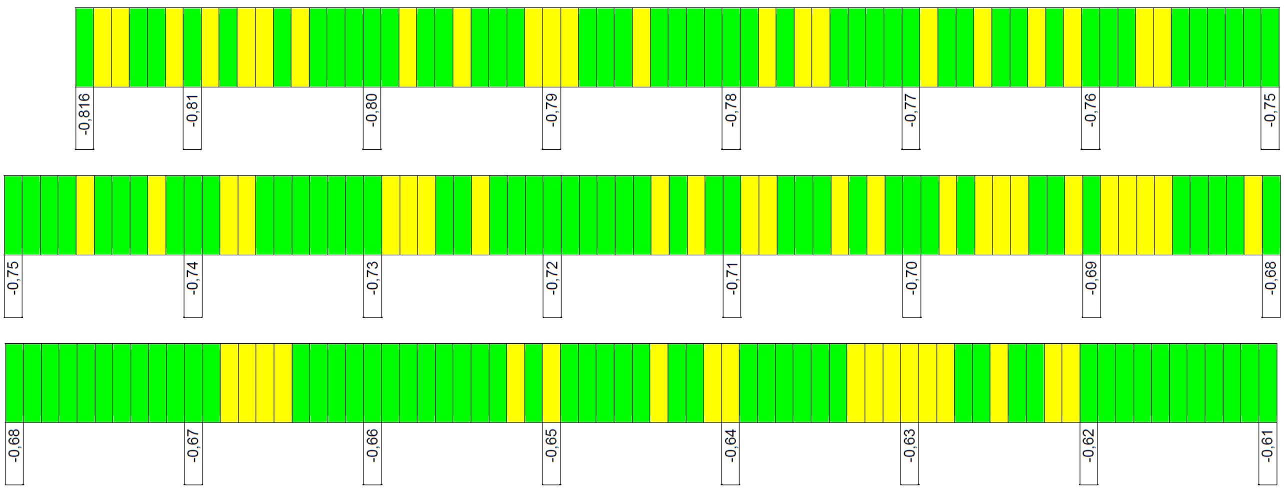
Bei der nächsten Stufe der Untersuchungen wurden Startwerte zwischen $x_0 = -1{,}0$ und $x_0 = +3{,}0$ mit $\Delta x = 0{,}02$ gewählt.



Im Folgenden wird noch beispielhaft ein einzelner Bereich mit der Schrittweite $\Delta x = 0{,}001$ untersucht – konkret ein Intervall rechts vom lokalen Maximum der Funktion *f*. Nach den bisherigen Untersuchungen hätte man vermuten können, dass Startwerte im gesamten Bereich zwischen dem lokalen Maximum und $x_0 \approx -0{,}64$ zu konvergenten Folgen führen.

Dieses Bild ändert sich jedoch bei näherem Hinsehen: Unvorhersehbar wechseln sich Konvergenz (grün) und Divergenz (gelb) ab. Man beachte dabei: Die Konvergenz- bzw. Divergenzeigenschaft bezieht sich jeweils nur auf die einzelnen angegebenen Stellen, die jeweils $\Delta x = 0{,}001$ voneinander entfernt liegen.

Wählt man einen kleineren Abstand zwischen den Startwerten, dann verändert sich das Bild in äußerst komplexer Weise.



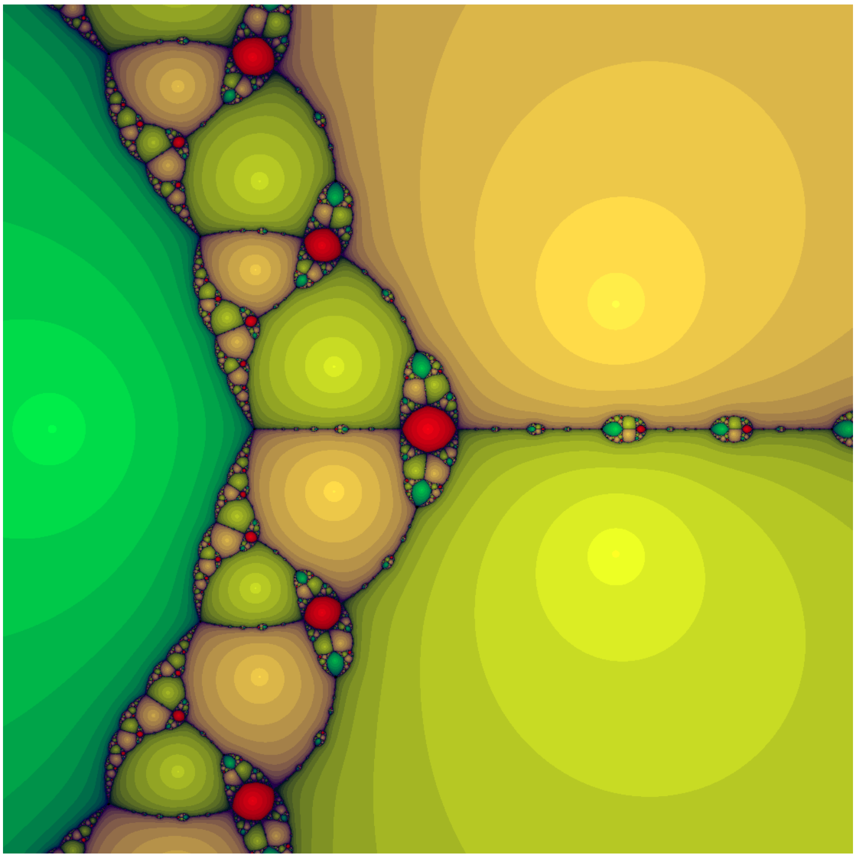
Newton-Fraktale

Das Newton-Verfahren kann auch für komplex-zahlige Argumente angewandt werden.

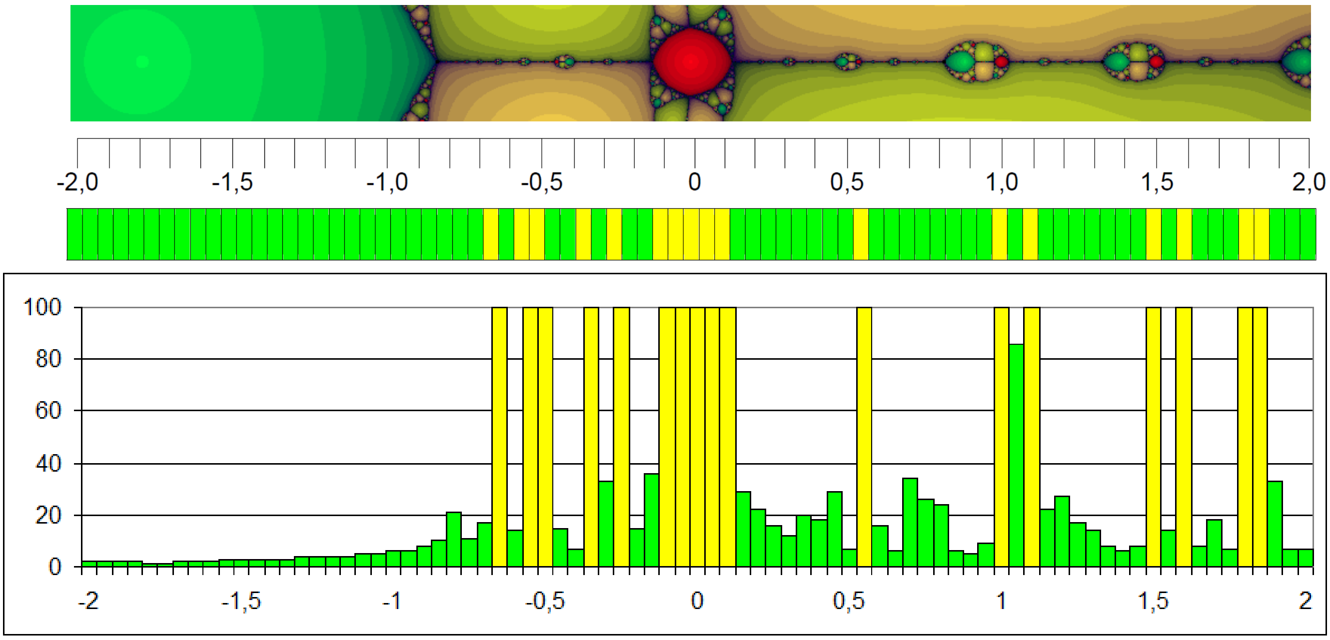
Im Wikipedia-Artikel *Newton fractal* findet man die folgende gemeinfreie Grafik von Henning Makholm mit dem Kommentar:

Newton fractal generated by the polynomial $z^3 - 2z + 2$*. Image area covering a square with side length 4 centered on 0. The corner pixels represent* $\pm 2 \pm 2i$ *exactly. Pixels are colored according to the number of iterations it takes to get closer than* $1e-3$ *either to a root or to 0.*

Im folgenden $[-2\,;+2] \times [-2i\,;+2i]$-Ausschnitt der Gauss'schen Zahlenebene findet man die beiden zueinander konjugiert-komplexen Nullstellen bei $z_1 \approx 0{,}88464 - 0{,}58974 \cdot i$ und $z_2 \approx 0{,}88464 + 0{,}58974 \cdot i$ sowie die reelle Nullstelle bei $z_3 = -1{,}7692923...$.

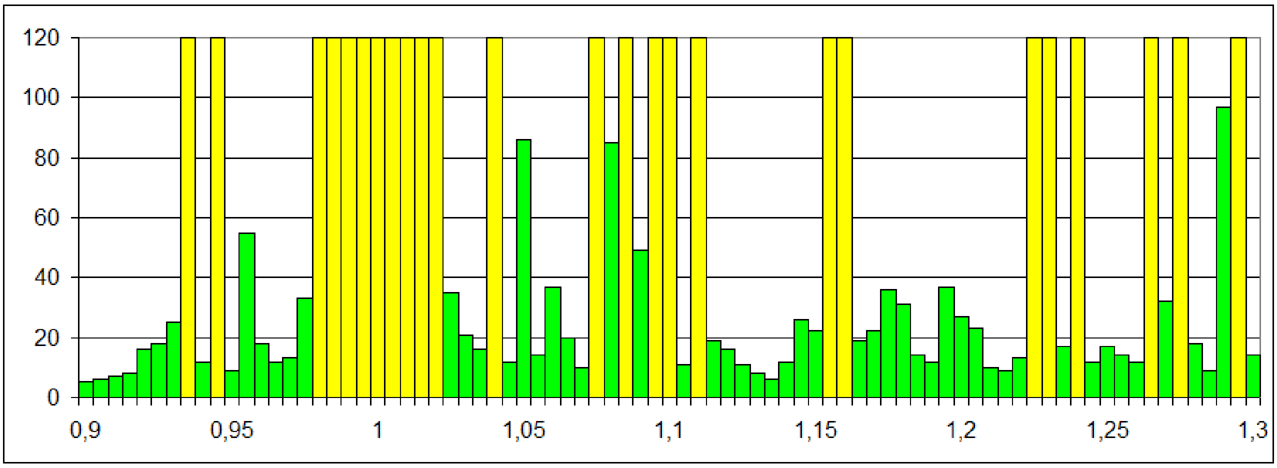


Die Grafik enthält in mittlerer Höhe auch die Darstellung des Konvergenzverhaltens für reellwertige Startwerte. Man erkennt insbesondere den Bereich links vom lokalen Maximum bei $x = -0{,}81649658...$, in dem das Verfahren stets konvergiert (grün), sowie das Intervall um den Ursprung, in dem alle Startwerte zu divergenten Folgen führen (rot), außerdem die komplexen Strukturen an anderen Stellen.



In der letzten Grafik ist festgehalten, wie viele Iterationsschritte erforderlich sind, bis ein Folgenglied sich vom o. a. Grenzwert um weniger als ein Tausendstel unterscheidet. Auch hier sind erhebliche Unterschiede erkennbar. Bemerkenswert ist die vergleichsweise große Anzahl an Iterationsschritten beim Startwert $x_0 = +1{,}05$; erst nach 86 Schritten ist die Bedingung $|x_n - (-1{,}76929...)| < \frac{1}{1000}$ erfüllt.

Die Veränderung der Schrittweite auf $\Delta x = 0{,}005$ zeigt weitere fraktale Strukturen.



Untersuchungen bei einer ganzrationalen Funktion 3. Grades mit drei reellen Nullstellen

Im Folgenden werde nun das Konvergenzverhalten bei der Funktion 3. Grades mit $f(x) = x^3 - 3x + 1$ untersucht. Das Beispiel wurde so gewählt, dass die beiden Extrempunkte der Funktion ganzzahlig sind: $H(-1\,|\,+3)$ und $T(+1\,|\,-1)$.

Der folgenden Wertetabelle kann man entnehmen, dass die drei Vorzeichenwechsel in den Intervallen $[-2\,;-1]$, $[0\,;+2]$ und $[+1\,;+2]$ stattfinden:

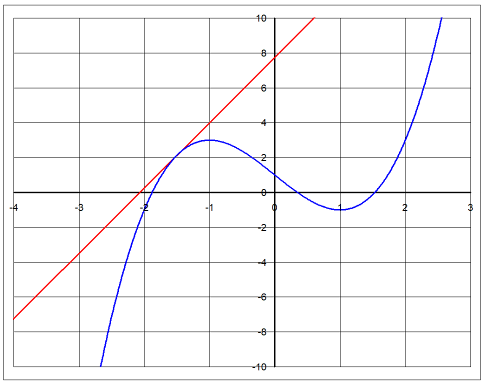
x	-2	-1	0	+1	+2
$f(x)$	-1	3	1	-1	3

Die drei reellen Nullstellen sind $a = -1{,}879385...$, $b = 0{,}347296...$ und $c = 1{,}532088...$.

Für die hier betrachtete Funktion *f* ergibt sich für das NEWTON-Verfahren die Rekursionsvorschrift

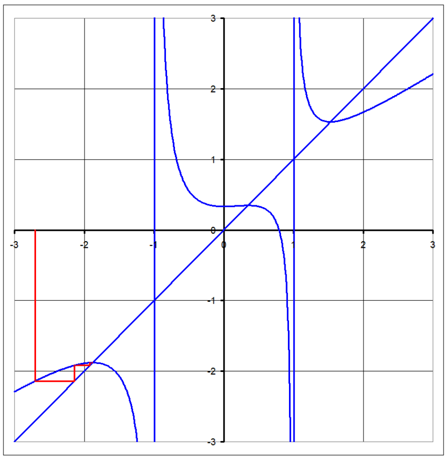
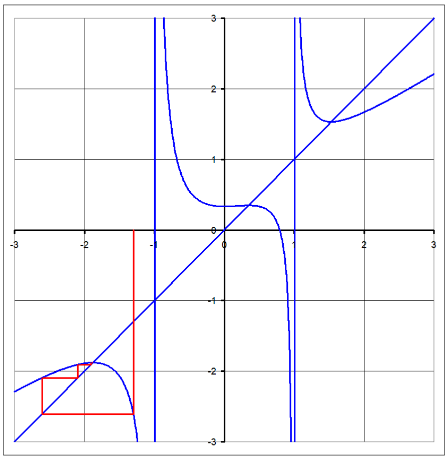
$$x_{n+1} = x_n - \frac{x_n^3 - 3x_n + 1}{3x_n^2 - 3} = \frac{3x_n^3 - 3x_n - x_n^3 + 3x_n - 1}{3x_n^2 - 3} = \frac{2x_n^3 - 1}{3x_n^2 - 3} =: g(x_n)$$

Die beiden Extremstellen der Funktion *f* sind die Polstellen der Funktion *g*.

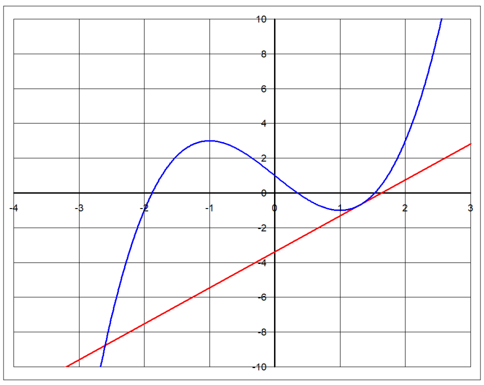


Tangenten durch einen Kurvenpunkt links vom Hochpunkt schneiden (wegen der Rechtskrümmung) die *x*-Achse links von *a*,

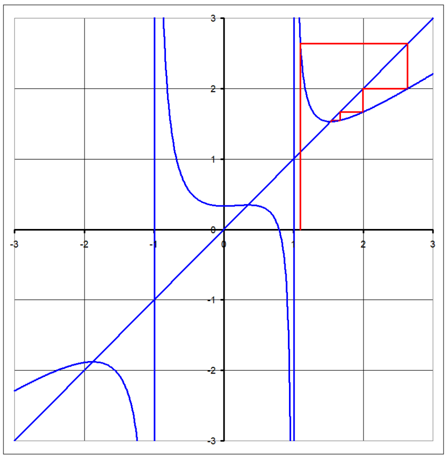
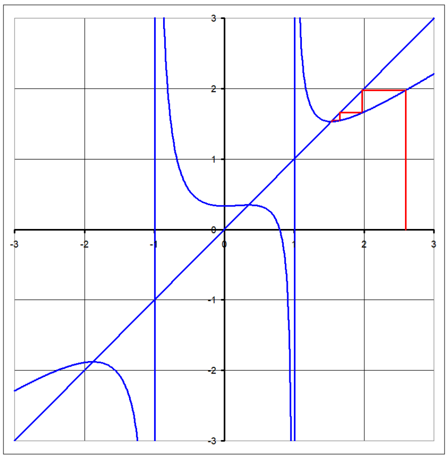
d. h., beginnt man die Iteration mit einem Startwert $x_0 < -1$, dann bleiben alle Folgenglieder der betreffenden Iteration auch unterhalb dieser Stelle, und die Folge konvergiert gegen die Nullstelle *a*. Die folgenden Beispiele zeigen die Spinnweb-Diagramme zu $x_0 = -2{,}7$ und $x_0 = -1{,}3$.

Tangenten durch einen Kurvenpunkt rechts vom Tiefpunkt schneiden (wegen der Linkskrümmung) die *x*-Achse rechts von *c*, d. h., beginnt man die Iteration mit einem Startwert $x_0 > +1$, dann bleiben alle Folgenglieder der betreffenden Iteration auch oberhalb dieser Stelle, und die Folge konvergiert gegen die Nullstelle *c*.



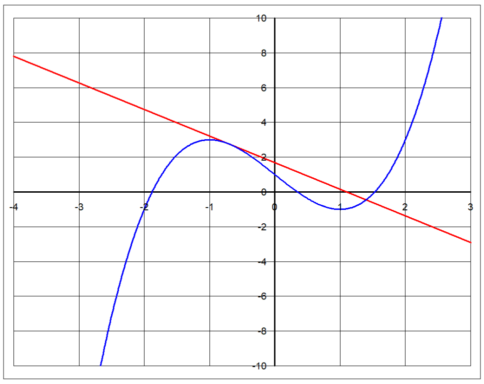
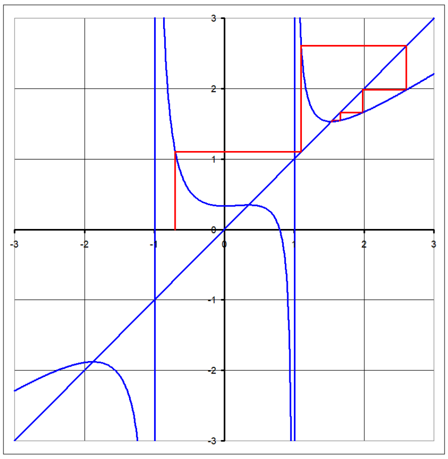
Die folgenden Beispiele zeigen die Spinnweb-Diagramme zu $x_0 = +1{,}1$ und $x_0 = +2{,}6$.

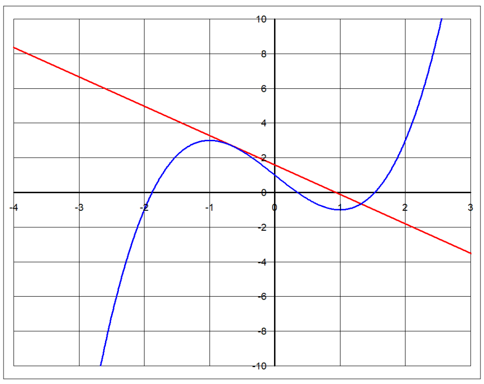
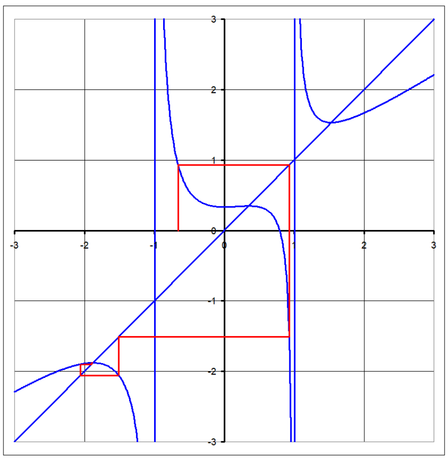
An den nächsten Beispielen kann man ablesen, dass bei der Wahl der Startwerte im Intervall $]-1;+1[$ bereits geringe Unterschiede zu unterschiedlichen Grenzwerten führen:

Die Iteration mit dem Startwert $x_0 = -0{,}7$ führt zur Nullstelle *c*; denn wie man der Abbildung mit dem Graphen von *f* und der Tangente durch den Startpunkt $(-0{,}7 \,|\, f(-0{,}7))$ entnehmen kann, schneidet die Tangente die x-Achse *rechts* vom Tiefpunkt von *f* und führt somit zu einer Iteration, die zur Nullstelle *c* führt.

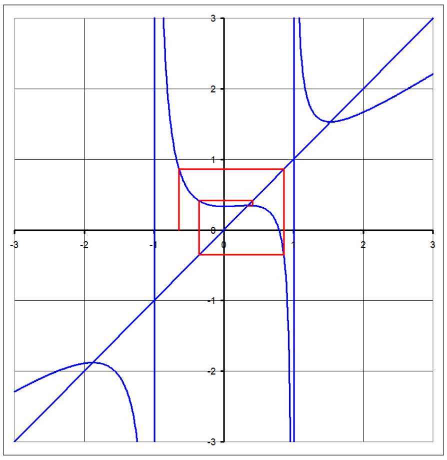
Dies kann auch aus dem zugehörigen Spinnweb-Diagramm rechts entnommen werden.

Auf der nächsten Abbildung mit der Tangente durch den Startpunkt $(-0{,}66 \,|\, f(-0{,}66))$ kann man so gerade noch erkennen, dass die Tangente an den Graphen von *f* die *x*-Achse *links* vom Tiefpunkt schneidet und daher zu einem anderen Konvergenzverhalten führt. Aus dem Spinnweb-Diagramm ist zu entnehmen, dass die Folge mit dem Startwert $x_0 = -0{,}66$ zur Nullstelle *a* führt.

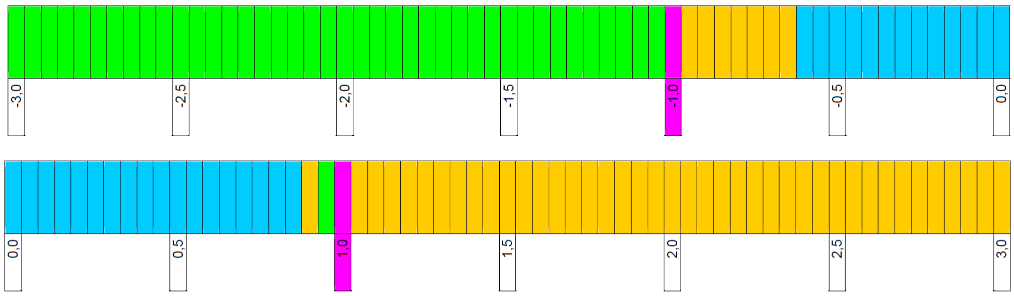
 

Auf eine Grafik mit Graphen der Funktion *f* und der Tangente wurde im nächsten Beispiel verzichtet, weil man den Unterschied zum vorangehenden Beispiel kaum erkennen kann; am Spinnweb-Diagramm kann man ablesen, dass die Folge mit dem Startwert $x_0 = -0{,}64$ zur Nullstelle *b* führt.



An den nächsten Grafiken kann man in einem ersten Überblick sehen, welche Startwerte zwischen $x_0 = -3{,}0$ und $x_0 = +3{,}0$ zu welchen Nullstellen-Grenzwerten führen. Dabei wurde für den Grenzwert *a* die Farbe *grün* gewählt, für den Grenzwert *c* die Farbe *goldgelb* und für den Grenzwert *b* die Farbe *himmelblau*. Die Polstellen $x = -1$ und $x = +1$ der Rekursionsfunktion sind durch die violette Färbung gekennzeichnet.

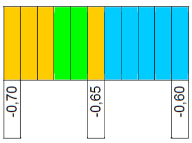
Im Einzelnen ergibt sich für Schrittweite $\Delta x = 0{,}05$:

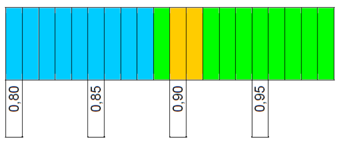


Durch Verfeinerung der Schrittweite entdeckt man weitere Einzelheiten, gegen welche Grenzwerte die Folgen in Abhängigkeit von den verschiedenen Startwerten konvergieren:

Schrittweite $\Delta x = 0{,}01$:

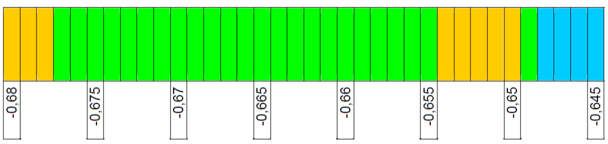
Untersuchungen im Intervall $[-0{,}70\,;-0{,}60]$ und im Intervall $[+0{,}80\,;+0{,}99]$

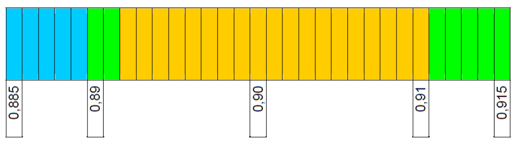




Schrittweite $\Delta x = 0{,}001$:

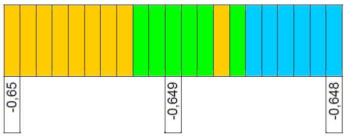
Untersuchungen im Intervall $[-0{,}680\,;-0{,}645]$ und im Intervall $[+0{,}885\,;+0{,}915]$

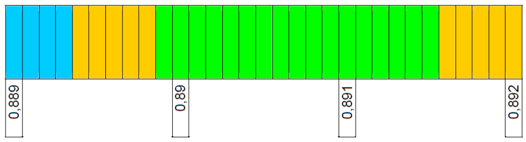




Schrittweite $\Delta x = 0{,}0001$:

Untersuchungen im Intervall $[-0{,}650\,;-0{,}648]$ und im Intervall $[+0{,}889\,;+0{,}892]$





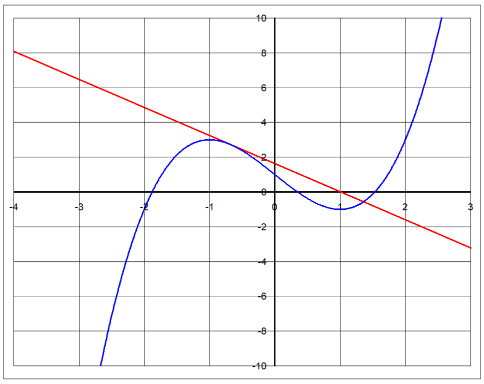
An den betrachteten Beispiel-Intervallen wird deutlich, dass bei Verfeinerung der Schrittweite vorher nicht sichtbare Unterschiede bzgl. der Grenzwerte aufgedeckt werden können.

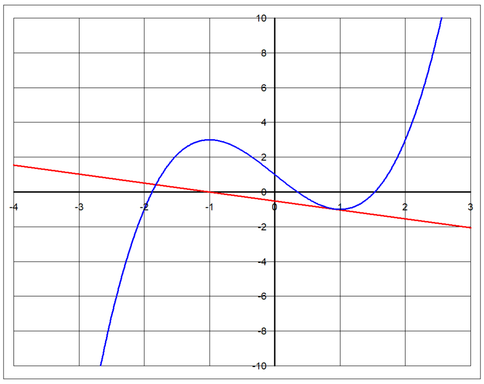
Untersuchung der Polstellen der Iterationsfunktion

Abschließend wollen wir uns mit den Polstellen der Iterationsfunktion *g* beschäftigen.

Die Startwerte $x_0 = -1$ und $x_0 = +1$ sind für die Iteration ja offensichtlich nicht geeignet, da die Tangenten parallel zur *x*-Achse verlaufen und daher keine Nullstelle besitzen.

Wandert man längs des Graphen der Funktion f und zeichnet jeweils die Tangente ein, dann kann man zumindest näherungsweise die beiden Tangenten finden, deren Nullstellen bei $x = -1$ bzw. $x = +1$ liegen. Die beiden zugehörigen Stellen sind daher als Startwerte für eine Iteration ebenfalls nicht geeignet.

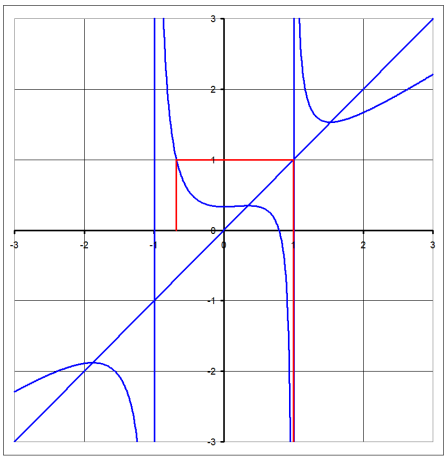


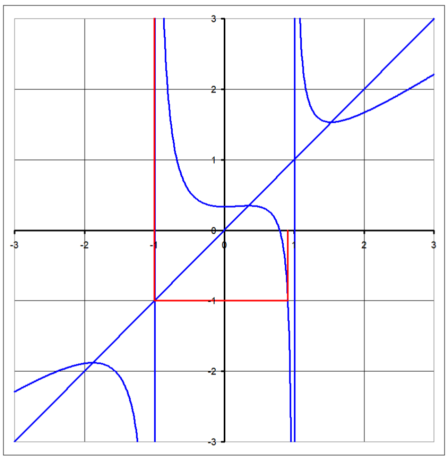


Bezogen auf das Iterationsverfahren werden also die beiden Stellen x_0 gesucht, für die gilt:

$$\frac{2x_0^3 - 1}{3x_0^2 - 3} = +1 = x_1 \text{ bzw. } \frac{2x_0^3 - 1}{3x_0^2 - 3} = -1 = x_1 .$$

Genauere Untersuchungen ergeben, dass diese Gleichungen die Lösung $x_0 = -0{,}6776506...$ bzw. die Lösung $x_0 = +0{,}9108200...$ haben.

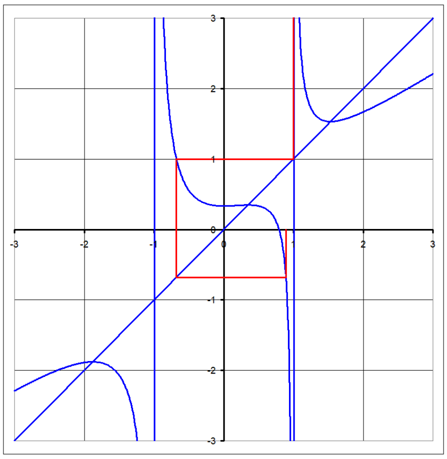


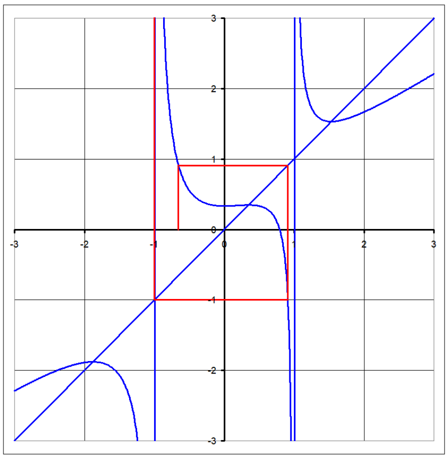


Doch auch diese beiden Startwerte können selbst auch als Folgenglied x_1 einer Iteration auftreten:

Die Gleichung $\frac{2x_0^3-1}{3x_0^2-3} = -0{,}6776506... = x_1$ hat die Lösung $x_0 = +0{,}8915165...$,

und die Gleichung $\frac{2x_0^3-1}{3x_0^2-3} = +0{,}9108200... = x_1$ hat die Lösung $x_0 = -0{,}6546982...$





Und die Gleichung $\frac{2x_0^3-1}{3x_0^2-3} = +0{,}8915165... = x_1$ hat die Lösung $x_0 = -0{,}6492179...$, und die Gleichung $\frac{2x_0^3-1}{3x_0^2-3} = -0{,}6546982... = x_1$ hat die Lösung $x_0 = +0{,}8898056...$

usw.

Im Prinzip lässt sich dieses Verfahren beliebig „rückwärts" fortsetzen: Tangenten an den angegebenen Stellen schneiden die *x*-Achse an der jeweils im benachbarten Feld rechts angegebenen Stelle. Darstellungen von Iterationen mit diesen Startwerten führen in einem Spinnweb-Diagramm also schließlich auf einen Punkt, der auf einer der beiden Polgeraden liegt.

…	+0,8893888...	–0,6492179...	+0,8915165...	–0,6776506...	+1
…	–0,6487223...	+0,8898056...	–0,6546982...	+0,9108200...	–1

Alle diese Stellen sind also als Startwerte für die Iteration nicht geeignet – bei allen besteht aber das grundsätzliche Problem, dass man sie nur mit einer gewissen numerischen Genauigkeit, also nicht exakt bestimmen kann.