Lipschitz-Stetigkeit

Hier werden wir ein wenig auf den Begriff der Lipschitz-Stetigkeit eingehen, diesen anschaulich erklären und überhaupt feststellen, wieso man einen weiteren Stetigkeitsbegriff braucht. Wir werden sehen, dass der Begriff genau zwischen der stetigen Differenzierbarkeit und der Stetigkeit steht. Am Ende werden wir noch ein passendes Beispiel mit expliziten Berechnungen bringen.

Zuerst erzähle ich mal etwas über die Benennung. Man sieht oft die Falschschreibungen von Lipschitz mit „Doppel p“, die man besser vermeiden sollte, denn der Name ehrt einen deutschen Mathematiker aus dem 19. Jahrhundert, nämlich Rudolf Lipschitz. Der Begriff hat sich auch in vielen anderen Sprachen gefestigt. Auf Englisch spricht man genauso von *Lipschitz continuous*, während man in Frankreich das Adjektiv *lipschitzienne* eingeführt hat. Wie man erwartet hat Rudolph Lipschitz diesen neuen Stetigkeitsbegriff entwickelt und vor allem in seinem *Lehrbuch der Analysis* aufbereitet.

Ich versuche den Begriff nun etwas moderner und anschaulicher vorzustellen. Wir nehmen dazu eine Funktion $f : \mathbb{R} \to \mathbb{R}$. Als gute Merkregel, sollte man sich nun folgende Eigenschaft der verschiedene Begriffe einprägen:

$$f \text{ stetig differenzierbar } \Longrightarrow f \text{ lokal Lipschitz-stetig } \Longrightarrow f \text{ stetig.}$$

Dieser neue Stetigkeitsbegriff soll also genau zwischen den zwei sehr bekannten Bezeichnungen aus der Analysis stehen. Sozusagen eine Abschwächung der stetigen Differenzierbarkeit. Was dieses „lokal“ an dieser Stelle zu suchen hat, werden wir gleich noch erklären. Es sei aber schon so viel gesagt, dass die sogenannte *lokale Lipschitz-Stetigkeit* der eigentlich interessante Begriff ist.

Nun haben wir aber lange genug um den heißen Brei herumgeredet. Wir müssen wie immer in der Mathematik eine Definition aufschreiben, um mit einem neuen Begriff umgehen zu können.

Definition der Lipschitz-Stetigkeit. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt *Lipschitz-stetig*, wenn es eine Konstante L gibt, welche

$$|f(x) - f(y)| \leq L\,|x - y| \tag{1}$$

für alle $x, y \in \mathbb{R}$ erfüllt.

Die Definition ist zwar kurz und knackig, aber in dieser Form noch nicht ganz selbsterklärend. An erster Stelle sollte man erwähnen, dass die Konstante L immer positiv oder Null sein muss, denn sonst kann die Ungleichung ja nie erfüllt sein. Weiterhin sieht man sofort, dass aus dieser Bedingung die Stetigkeit der Funktion f an jeder Stelle folgt, denn:

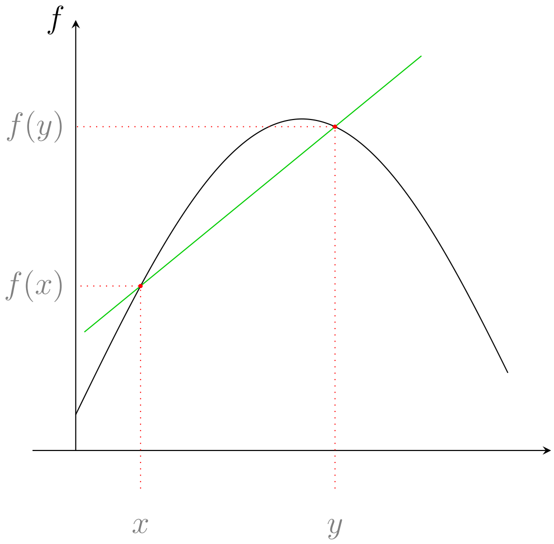
$$|f(x) - f(y)| \leq L\,|x - y| \xrightarrow{x \to y} 0\ .$$

Oder anders gesagt: Wenn die Zahlen x, y immer näher zusammenrücken, dann müssen dies auch die Funktionswerte tun. Dies ist nichts anderes als die Stetigkeit.

Was hat dies nun aber mit der Differenzierbarkeit zu tun? Um dies zu erkennen, muss man die Definition nur ganz ganz wenig umschreiben. Wir dividieren auf beiden der Gleichung (1) durch $|x - y|$, natürlich unter der Annahme, dass nun $x \neq y$ gilt. Dann ist die Lipschitz-Stetigkeit gleichbedeutend mit der Eigenschaft

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq L\,, \qquad \text{für alle } x \neq y\,. \tag{2}$$

Nun steht auf der linken Seite ein wohlbekannter Differenzenquotient, bzw. eine Sekantensteigung.



Nun können wir auch die zu Beginn festgestellte Folgerung, nämlich, dass aus der Differenzierbarkeit die Lipschitz-Stetigkeit folgt, anschaulich verstehen. Dazu ein sehr allgemeines Beispiel.

Differenzierbarkeit und Lipschitz-Stetigkeit. Betrachtet man eine differenzierbare Funktion $f : I \to \mathbb{R}$ auf einem offenen Intervall I, so gilt bekanntlich der Mittelwertsatz:

$$f'(\xi_{x,y}) = \frac{f(x) - f(y)}{x - y} \ .$$

Also für zwei beliebige Punkte x, y gibt es eine Zwischenstelle $\xi_{x,y}$ derart, dass die Sekantensteigung von oben, durch die Ableitung an dieser Zwischenstelle gegeben ist. Im obigen Bild lässt sich diese Zwischenstelle leicht erraten. Wenn man nun alle möglichen Werte von x, y durchläuft, so durchlaufen die möglichen Zwischenstellen die gesamte Ableitung, sodass (2) zu

$$|f'(z)| \leq L \,, \qquad \text{für alle } z \in I$$

wird. Das heißt aber nichts anderes, dass wir L durch

$$L := \sup_{z \in I} |f'(z)|$$

wählen können und die differenzierbare Funktion genau dann Lipschitz-stetig ist, wenn das so definierte L kleiner Unendlich ist.

Das obige Beispiel ist auch das, was typischerweise in Anwendungen vorkommt. Man muss demnach nur die Ableitung berechnen und schauen, dass diese beschränkt bleibt. Wir fassen dies in der folgenden Merkregel zusammen

Merkregel. Eine Lipschitz-stetige Funktion ist eine stetige Funktion, deren Steigungen beschränkt bleiben. Dabei ist es unerheblich, ob wir diese Steigung als Ableitung an einem Punkt messen, oder als Sekantensteigung an zwei Punkten.

Zum Abschluss beschreiben wir noch kurz die lokale Lipschitz-Stetigkeit in einer kurzen Merkregel. Von einer *lokal* Lipschitz-stetigen Funktion $f : \mathbb{R} \to \mathbb{R}$ spricht man, wenn man die Lipschitz-Bedingung (1) nur auf beschränkten Intervallen fordert, d. h. es ist unerheblich, was „im Unendlichen" geschieht. Es geht hier nur um das lokale[1] Verhalten.

Merkregel. Eine lokal Lipschitz-stetige Funktion $f : \mathbb{R} \to \mathbb{R}$ ist eine stetige Funktion, deren Steigungen auf allen beschränkten Intervallen beschränkt bleiben. Dabei ist es unerheblich, ob wir diese Steigung als Ableitung an einem Punkt messen, oder als Sekantensteigung an zwei Punkten.

Nach dieser Merkegel muss die Funktion also keineswegs differenzierbar sein, aber kann trotzdem lokal Lipschitz-stetig sein. Wenn die Funktion aber nun doch differenzierbar ist, d. h. es gibt eine Ableitung die auf dem gesamten betrachteten Bereich existiert, kann sie dann womöglich doch nicht lokal Lipschitz-stetig sein?

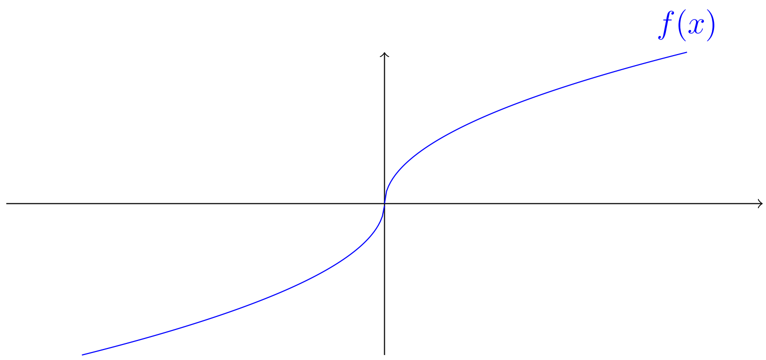
Die Frage ist nicht leicht zu beantworten, denn alle typischen Gegenbeispiele von nicht lokal Lipschitz-stetigen Funktionen sind entweder noch nicht einmal

[1] Betrachtet man Funktionen, die nicht auf der gesamten reellen Zahlengerade definiert sind, so muss man den Begriff genauer definieren: Eine Funktion $f : D \to \mathbb{R}$ heißt dann *lokal Lipschitz-stetig*, wenn es für jeden Punkt $x \in D$ eine Umgebung $U \subseteq D$ gibt, auf der die Lipschitz-Bedingung (1) erfüllt ist.

stetig oder es sind angenehme Funktionen, die sogar bis auf einen Punkt differenzierbar sind. Dazu sollte ich vielleicht wirklich das typische Beispiel hier präsentieren:

$$f(x) = \begin{cases} \sqrt{x} & , \quad x \geq 0 \\ -\sqrt{-x} & , \quad x < 0 \end{cases}$$

Es hat den folgenden Funktionsgraphen:



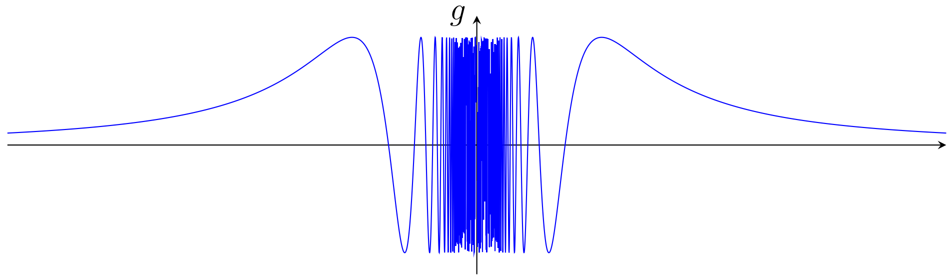
Der Graph bekommt, wenn man sich der Null nähert, eine unendliche Steigung. Die Funktion ist also bis auf die Null differenzierbar, aber da die Steigungen, um die Null nicht beschränkt bleiben, kann die Funktion nicht lokal Lipschitz-stetig sein.

Das Problem an diesem Beispiel ist, dass f ja überhaupt nicht *überall* differenzierbar ist, sodass unsere Frage weiterhin offen bleibt. Wir suchen also eine Funktion, die wirklich auf ganz $\mathbb{R}$ differenzierbar ist, aber trotzdem beliebig hohe Steigungen um die Null aufweisen kann.

Beliebige Steigungen erhalten wir, indem wir die Funktion um den Ursprung immer mehr Schwingungen ausführen lassen, z.B. mit einer Sinusfunktion:

$$g(x) = \sin\left(\frac{1}{x^2}\right)$$

mit dem folgenden Kurvenverlauf:

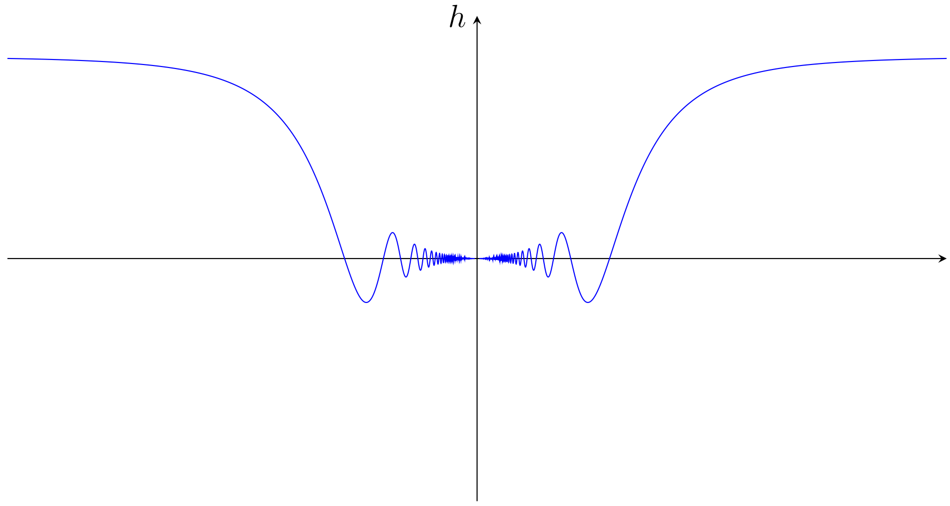


Da die Funktion immer mehr Schwingungen durchführt, wird es ziemlich dicht um die Null herum, sodass man dort grafisch kaum mehr etwas erkennen kann. Aber man erkennt zumindest, dass die Funktion immer größere Steigungen bekommt. Leider hat nun auch diese Funktion den Nachteil, dass sie an der Null nicht differenzierbar ist, ja der Differentialquotient existiert nicht mal im uneigentlichen Sinne, da die Steigung zwischen negativen und positiven Werten hin- und herpendelt. Und noch viel schlimmer: Die Funktion ist nicht mal in der Null definiert, geschweige denn stetig fortsetzbar.

Wir müssen nun also an dieser Stelle etwas nachdenken. Aber es hatte ja seinen Grund, warum wir diese Funktionen genommen haben. Die Funktion muss gegen die Null nur passend abfallen, damit der Differentialquotient existiert. Wir nehmen letztendlich die folgende Funktion $h : \mathbb{R} \to \mathbb{R}$, die durch

$$h(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases}$$

definiert ist. Auch diese besitzt wieder das Schwingungsverhalten, d. h. es werden immer größere Steigungen angenommen. Dies muss man natürlich an dieser Stelle nachrechnen, wobei man dies durch die Ableitung mit Hilfe der Produkt- und Kettenregel erhält. Dazu auch wieder der passende Graph:



Wir wissen nun also, dass die Funktion außerhalb der Null differenzierbar ist und dass die maximale Steigung in der Nähe der Null immer größere Werte bekommt. Das heißt, die Funktion ist auf jeden Fall *nicht* lokal Lipschitz-stetig. Wir müssen nun nur noch zeigen, dass die Funktion auch für $x = 0$ differenzierbar ist, d. h. einen endlichen Wert für den Differentialquotienten besitzt. Diesen können wir nur aufschreiben:

$$\frac{h(x) - h(0)}{x - 0} = \frac{x^2 \sin(x^{-2})}{x} = x \sin(x^{-2}) \xrightarrow{x \to 0} 0 \ .$$

Wunderbar! Die Funktion ist auch also auf ganz $\mathbb{R}$ differenzierbar mit der Steigung 0 im Ursprung, was übrigens durch die einhüllende Funktion x^2 verursacht wird.

Wir haben also, im Nachhinein relativ leicht, eine differenzierbare Funktion aufschreiben können, die aber nicht lokal Lipschitz-stetig ist.