

Manual zur Auswertungssoftware ‘Levigator’ von EML-Experimenten

Stephan Rietzler

Annahmen und Auswertungsideen

Aufgrund theoretischer Überlegungen sollten die Signale (z.B. R_RadiusX), welche die Oszillation des flüssigen Metalltropfens beschreiben, folgende Signalform haben: $y = a \cdot e^{-b(T) \cdot t} \cdot \sin(2\pi \cdot f(T) \cdot t)$, wobei t =Beobachtungszeit, d.h. $t=0$ bezeichnet den Start der Aufnahme und T =Temperatur der Probe.

Die für die Materialeigenschaft zentrale Größe Temperatur wird kontinuierlich gemessen, d.h. man kennt $T = T(t)$ und man bekommt damit: $y = a \cdot e^{-b(T) \cdot t} \cdot \sin(2\pi \cdot f(T) \cdot t)$

Die Frequenz $f(T)$ und die Dämpfung $b(T)$ stehen in folgendem Zusammenhang mit den Materialeigenschaften Viskosität $\eta(T)$ und Oberflächenspannung $\sigma(T)$:

$$\begin{aligned} - \eta(T) &= \frac{3}{20\pi} \cdot \frac{m}{r} \cdot b(T) \\ - \sigma(T) &= \frac{3}{8} \pi m f(T)^2 \end{aligned}$$

Aus den optischen Daten werden nun Signale extrahiert welche die Oszillation der flüssigen Metallegierung beschreiben. (Details: Tevi-Software)

Aufgabe ist es nun, aus diesen Signalen ‘gute’ Schätzungen für die Zeit- bzw. Temperatur- abhängige Frequenz und Dämpfung des Signals zu bekommen.

Hierzu unterteilt man die Beobachtungszeit in Zeitfenster gleicher (!) Länge und unterstellt in einem solchen Zeitfenster folgende Signal-Form: $y = a \cdot e^{-b \cdot t} \cdot \sin(2\pi \cdot f \cdot t)$, d.h. konstante Dämpfung und Frequenz. Wobei das ‘Zeitfenster’ zur bestimmung der Dämpfung mehrere Zeitfenster zur bestimmung der Frequenz umfasst. Zur Bestimmung der Dämpfung kann man also nicht mehr diese ‘ideale’ Signalform unterstellen.

Die Schätzung der Frequenz und Dämpfung aus den Rohsignalen beruht auf folgendem Ansatz: In einem gegebenem Zeitfenster wird das Signal Fourier transformiert, und anhand der (geglätteten) Absolut-Beträge der Fourier-Koeffizienten die dominante Frequenz f_D mit zugehörigem Absolut-Betrag $A(f_D)$ des entsprechenden Fourier-Koeffizienten ermittelt.

Da die Zeitfenster alle gleich (!) lang sind ist $A(f_D)$ ein Maß für die Signal-Amplituden der dominanten Frequenz f_D in jedem Zeitfenster mit dem die Dämpfung des dominanten Signals über mehrere Zeitfenster ermittelt werden kann. D.h. die Dämpfung $b(T)$ wird ermittelt indem eine Exponential-Funktion $a \cdot e^{-b \cdot t}$ an $A(f_D(T(t)), T(t))$ gefittet wird. Hierbei muß man annehmen das die Temperatur über den Bereich über den gefittet wird konstant ist. Dies ist aufgrund der schnellen Abkühlung des Tropfens sicher nicht gegeben. Die Angegebene Dämpfung (und damit Viskosität) ist also bestenfalls ein Mittelwert über den entsprechenden Temperaturbereich.

Zur Bestimmung von $\eta(T)$ wird *generell* ein Zeitfenster gewählt welches durch eine Heizpuls begrenzt ist. D.h. bei zwei Heizpulsen bekommt man drei Zeitfenster in denen eine Viskosität bestimmt werden kann. Die zur jeweiligen Größe angegebene Temperatur entspricht dabei immer dem Mittelwert der Temperatur in dem verwendeten Zeitfenster.

Andere Methoden zur Schätzung der Signaldämpfung sind möglich und werden diskutiert.

Zur Bestimmung der Oberflächenspannung benötigt man nur die Frequenz der Schwingung. Diese Frequenz ermittelt man in jedem gewählten Zeitfenster als die Frequenz mit der ‘maximalen’ Fourier-Koeffizienten-Amplitude. Wobei diese geglättet und etwas aufbereitet werden. Aus dieser Frequenz bestimmt man dann zu jedem Zeitpunkt (Mittelpunkt des gewählten Zeitfensters) die Oberflächenspannung $\sigma(T(t))$. Die Temperaturabhängigkeit $\sigma(T)$ wird nun durch einen linearen Fit an die in den einzelnen Zeitfenstern ermittelten Oberflächenspannungen $\sigma(T)$ bestimmt.

Generelle Bemerkungen

Zur Auswahl der Signale: ‘Gemessen’ werden zunächst vier Größen. Nämlich die Haupttradien zweier um 45° gegeneinander gedrehter Ellipsen. Offen ist noch wie diese Größen von der Tevi-Software ermittelt werden. *Würde eine Ellipse konstanter Fläche gefittet (was ein Versuch wäre die Randbedingung: Konstantes Volumen zu berücksichtigen) so müßten die Signale x und y exakt 180° phasenverschoben sein!* Eine genaue Kenntniss wie die Signale von der Tevi-Software ermittelt werden wäre sicher hilfreich.

Signale wie z.B. Radius-Summe und -Differenz bieten aufgrund physikalischer Überlegungen gewisse Vorteile. Die Signale der Radien der beiden Hauptachsen sollten theoretisch 180° phasenverschoben sein. Durch Differenzbildung sollte sich dadurch ein deutlich verstärktes Signal der Schwingung ergeben. Zudem hat die Differenzbildung den Vorteil das sich Radienvergrößerungen und -verkleinerungen durch Bewegung der Kugel auf die Kamera zu aufheben. Statistisch haben sie aber den Nachteil das sie als Summe von Meßsignalen ein höheres Rauschen aufweisen. Das erwartete bessere signal-rausch-verhältnis dieser signale ist im Regelfall nicht zu beobachten. Aus diesem Grund werden default-mäßig **nur die Signale Radius X und Y ausgewertet**. Die Radien der gedrehten Ellipse werden ebenfalls defaultmäßig nicht ausgewertet, da nicht wirklich klar ist wieviel unabhängige Informationen diese Signale zusätzlich liefern.

Heuristische Analysen legen nahe das sich die Endergebnisse im Rahmen der Messgenauigkeit nicht ändern wenn man diese Signale hinzunimmt.

Aus Gründen der Effizienz werden deswegen nur die erwähnten Signale untersucht.

Die aus einem Datensatz (d.h. eine Kamera) geschätzten Viskositäten zu einer Temperatur werden mit einem Standardfehler se als Maß für die Messgenauigkeit angegeben. Umfasst das Intervall Messwert $\pm 2 \cdot se$ die Null, so wird dieser Messwert aufgrund der hohen Messungenauigkeit nicht in die Auswertung aufgenommen.

Vergleicht man die empirischen Daten mit simulierten Signalen so erweist es sich als nicht ganz einfach offensichtlich vorhandene Rauschen der Signale zu reproduzieren. Es scheint als ob das Rauschen amplitudenabhängig ist. D.h. bei großen Amplituden scheint es als ob die Signale stärker verrauscht sind. Die könnte durch die instabile Signalgenerierung durch Tevi verursacht sein. Bei großen Amplituden, d.h. am Anfang der Auswertung oszilliert der Tropfen noch sehr stark, wodurch er immer wieder in die Bildberandung läuft und der Algorithmus dann die Signale total verfälscht. Ich denke dieses Problem könnte man durch einen stabilen Fit einer Ellipse und bessere Kantendetektion beheben.

Auffälligkeiten

1. Es scheint niederfrequentes additives Rauschen zu dominieren. D.h. die Unterstellung daß das additive Rauschen weiß ist scheint nicht zuzutreffen.

Schätzung der Dämpfung aus den Rohsignalen

Zur Bestimmung der Dämpfung aus den gemessenen Signalen werden drei Ansätze vorgestellt und bzgl. ihrer Eigenschaften in Simulationen verglichen.

Zur Wahl des Simulations-Modells: Oft sind im Signal Modulationen erkennbar, so daß man nicht von dem erwarteten idealen Signal $ae^{-bt}\sin(2\pi ft)$ ausgehen kann. Außerdem erwartet man eine leichte Drift der Frequenz. Zur Untersuchung der Qualität der verschiedenen Ansätze zur Bestimmung der Dämpfungszeitkonstanten wird also ein Signal mit dem deterministischen Teil $ae^{-bt}\sin(2\pi \cdot (f + t/10) \cdot t)\sin(2\pi 0.5t)$ verwendet.

Um das Signal-Rauschen zu modellieren ist folgendes zu berücksichtigen: Aufgrund des Detektions-Algorithmuses von Tevi ist davon auszugehen, dass das Rauschen von der Amplitude abhängt. Konkret: Bei großen Amplituden am Anfang beobachtet man ein stärkeres Signal-Rauschen, da die Probe insgesamt unruhiger ist und oft auch in die Bild-Berandung ‘läuft’ was sich in einem starken Signal-Rauschen bemerkbar

macht. Aus diesem Grund wird ein multiplikativer und additiver Noise-Term in das Simulations-Modell aufgenommen. Als Ansatz für die Signal-Simulation wird deswegen verwendet:

$$y(t) = \left(ae^{-bt} \sin(2\pi \cdot (f + t/10) \cdot t + \pi/2) \sin(2\pi 0.5t + \pi/2) \right) \cdot \epsilon_*(t) + \epsilon_+(t)$$

$\epsilon_*(t)$ und $\epsilon_+(t)$ sind iid Normalverteilt. Dieses Signal wird mit 200Hz abgetastet.

Zur Veranschaulichung und zum Vergleich der Ansätze werden beispielhaft folgende Parameter verwendet:

$$y(t) = \left(20e^{-0.7t} \sin(2\pi \cdot (30 + t/10) \cdot t + \pi/2) \sin(2\pi 0.5t + \pi/2) \right) \cdot \epsilon_*(t) + \epsilon_+(t)$$

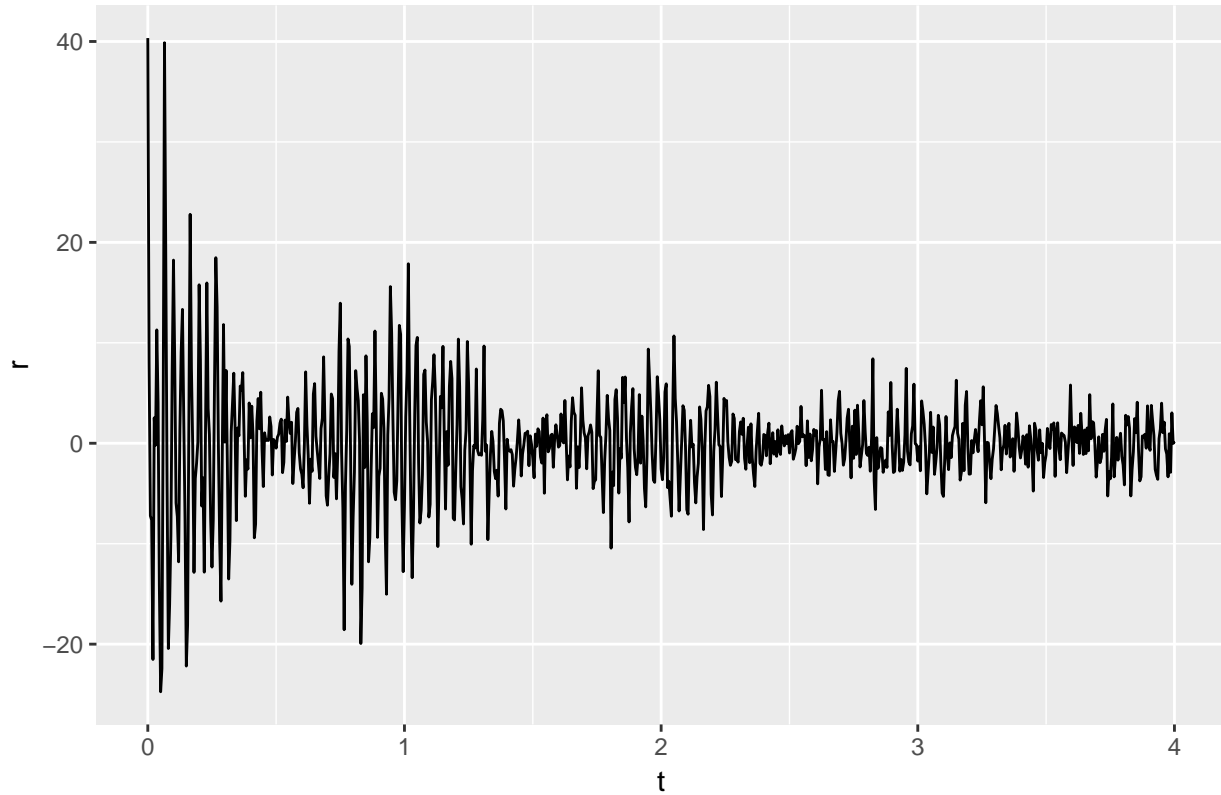
$$y(t) = \left(20e^{-0.5t} \sin(2\pi \cdot (32 + t/10) \cdot t + \pi/2) \sin(2\pi 0.5t + \pi/2) \right) \cdot \epsilon_*(t) + \epsilon_+(t)$$

$$y(t) = \left(20e^{-0.3t} \sin(2\pi \cdot (34 + t/10) \cdot t + \pi/2) \sin(2\pi 0.5t + \pi/2) \right) \cdot \epsilon_*(t) + \epsilon_+(t)$$

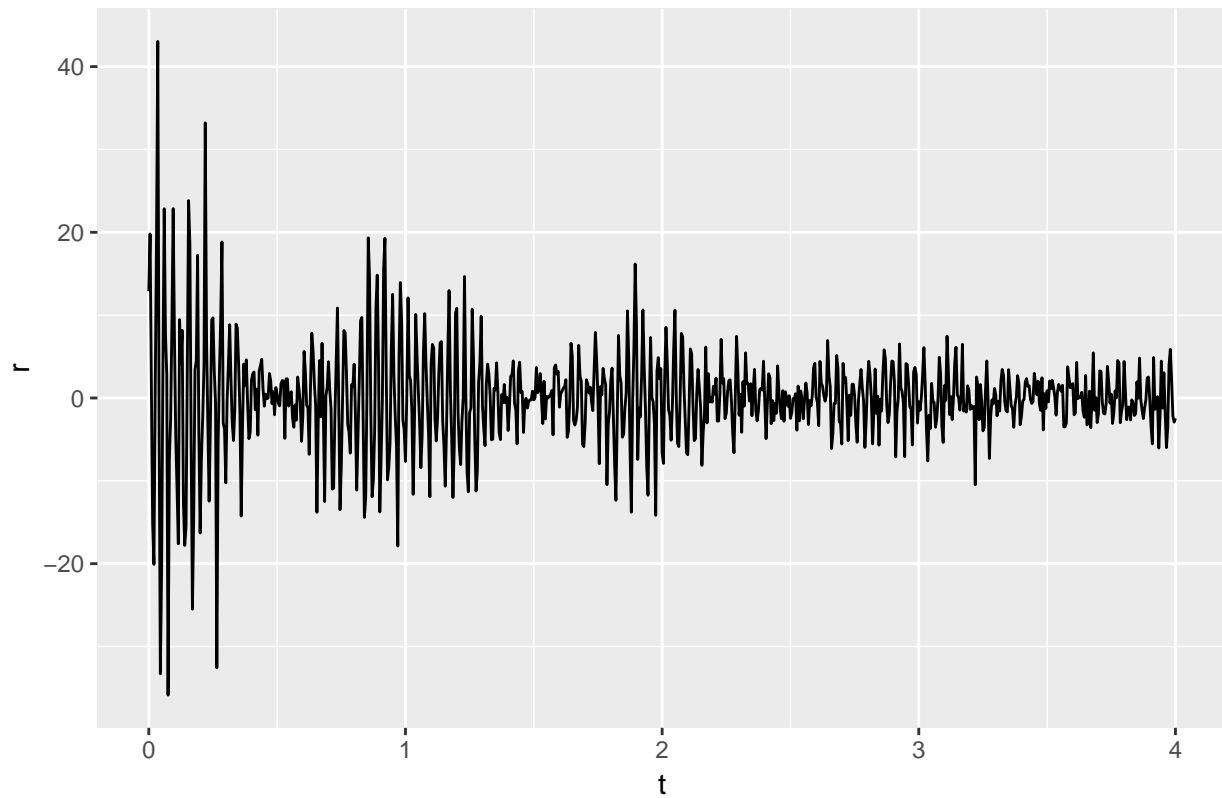
mit $\epsilon_*(t)$ iid $N(1, 0.5)$ und $\epsilon_+(t)$ iid $N(0, 2)$

Das Signal wird im Intervall $t \in [0, 4]$ mit einer samplerate = 200Hz generiert und ausgewertet. Folgende Plots zeigen jeweils eine Realisierung des Signals mit den unterschiedlichen Parametern:

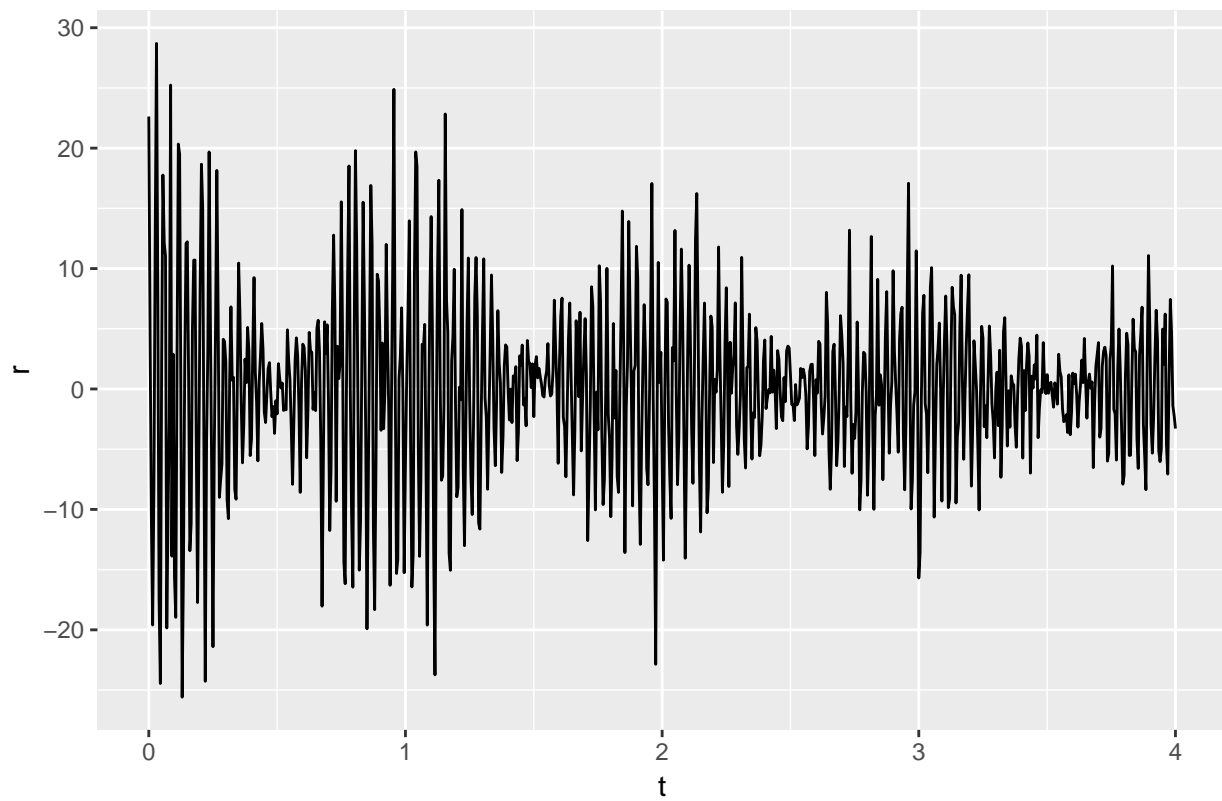
damping-const.: 0.7 1/s, f= 30 Hz



damping-const.: 0.5 1/s, $f = 32$ Hz



damping-const.: 0.3 1/s, $f = 34$ Hz



Ansatz: Signal-Transformation und linearer Fit

Das Rohsignal y wird quadriert und log-transformiert:

$$\tilde{y}(t) = \log(y(t)^2) \approx \log(a^2) - 2bt + \log(\sin(\omega t)^2) + \log(\epsilon_*^2)$$

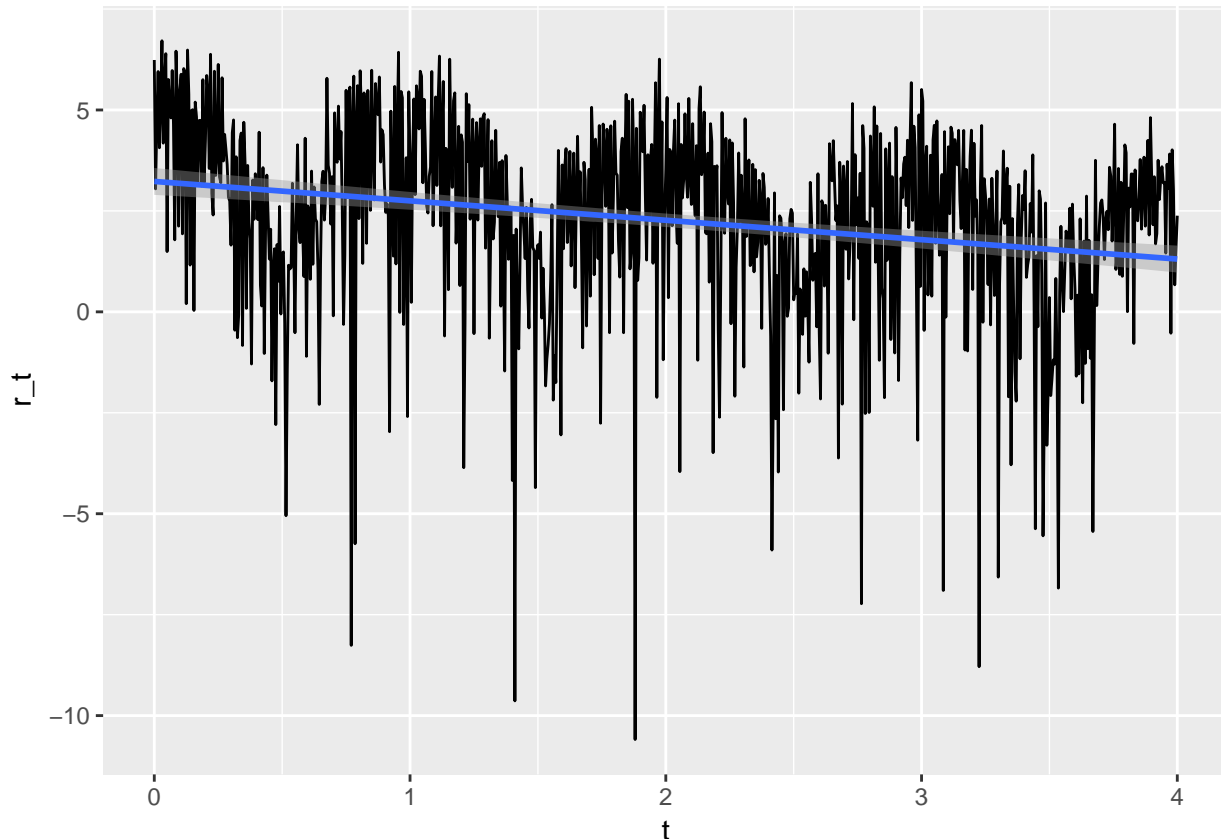
An dieses Signal wird eine Gerade gefittet. Eine Schätzung der Dämpfung b bekommt man dann aus der Schätzung der Geradensteigung m durch: $b = -\frac{m}{2}$. Der Standardfehler se_b (Messunsicherheit) von b ergibt sich dann aus dem geschätzten Standardfehler se_m von m : $se_b = \frac{1}{2} \cdot se_m$.

Die korrekte Schätzung von se_m beruht dabei auf der Annahme dass das signal von folgender Form ist: $y(t) = a + mt + \epsilon(t)$, wobei $\epsilon(t)$ iid normalverteilt ist.

Aufgrund der Transformation sieht man: $\epsilon(t) = \log(\sin(\omega t)^2) + \tilde{\epsilon}(t)$, d.h. $\epsilon(t)$ ist bei diesem Ansatz offensichtlich sehr stark (positiv) autokorreliert (dies ist an den nachfolgenden Plots auch offensichtlich erkennbar) wodurch se_m und damit se_b nur bedingt aussagekräftig sind!

Beispielhafte Analyse mit $b=0.7$ und $f= 30\text{Hz}$

Graphische Veranschaulichung der Methode



Numerische Analyse

```
##  
## Call:  
## lm(formula = r_t ~ t, data = sample_signals %>% mutate(r = s0,  
##       r_t = log(r^2)))  
##  
## Residuals:
```

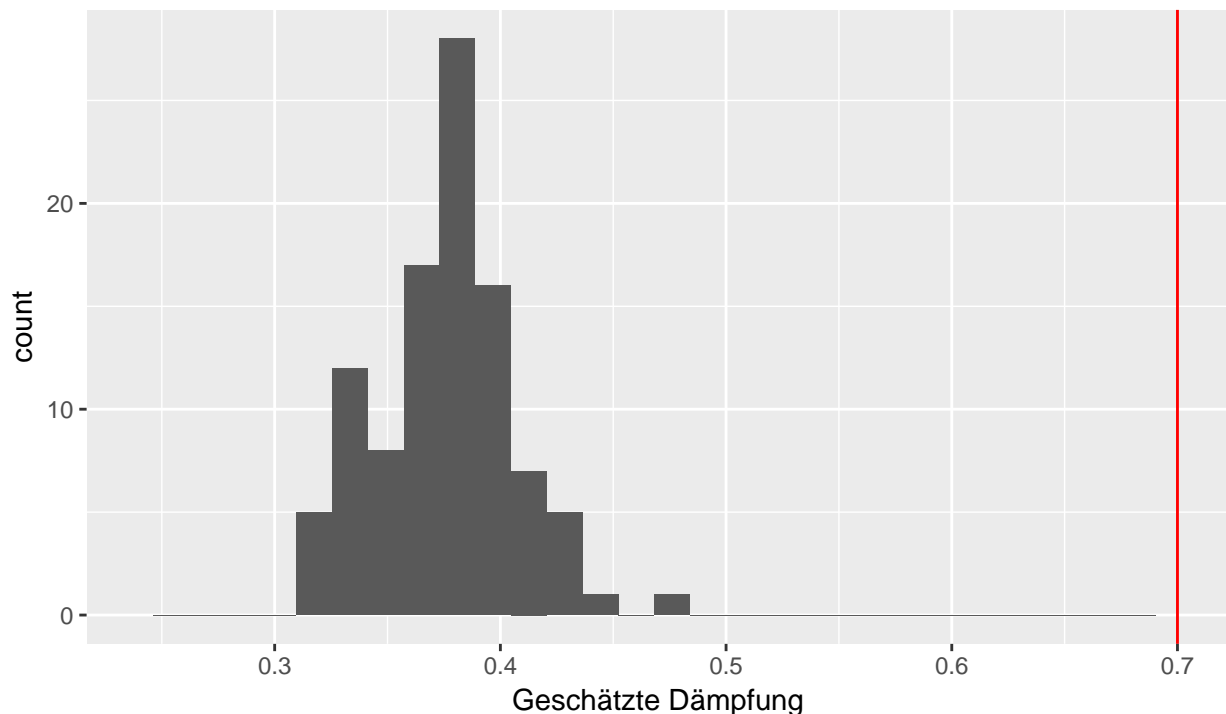
```
##      Min      1Q   Median      3Q      Max
## -12.9171 -1.1844  0.4317   1.7127  3.9779
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.2300     0.1696  19.050 < 2e-16 ***
## t            -0.4806     0.0734  -6.548 1.04e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.402 on 799 degrees of freedom
## Multiple R-squared:  0.05093,    Adjusted R-squared:  0.04974
## F-statistic: 42.88 on 1 and 799 DF,  p-value: 1.042e-10
```

Für dieses Beispiel liefert der Ansatz ein Ergebnis welches um den Faktor 2 zu klein ist. An der graphischen Darstellung erkennt man den Grund: Ab ca. 2.5s wird das Signal stark durch das additive Rauschen dominiert. Durch die Transformation nivelliert sich das Signal auf den Erwartungswert des log-transformierten quadratischen Rauschens und fällt nicht weiter linear ab. Aus diesem Grund wird die Gerade zu Ende 'angehoben' und dadurch die Steigung und damit die Dämpfung zu klein geschätzt. Dieses Problem wird umso schwerwiegender je größer die 'wahre' Dämpfungskonstante und je größer der Beobachtungszeitraum ist.

Im folgenden sind die geschätzten Ergebnisse der Dämpfung für 100 Realisierungen in einem Histogramm dargestellt:

Schätzung von μ mit durch linearen fit an quadriertes log-transformiertes sigr

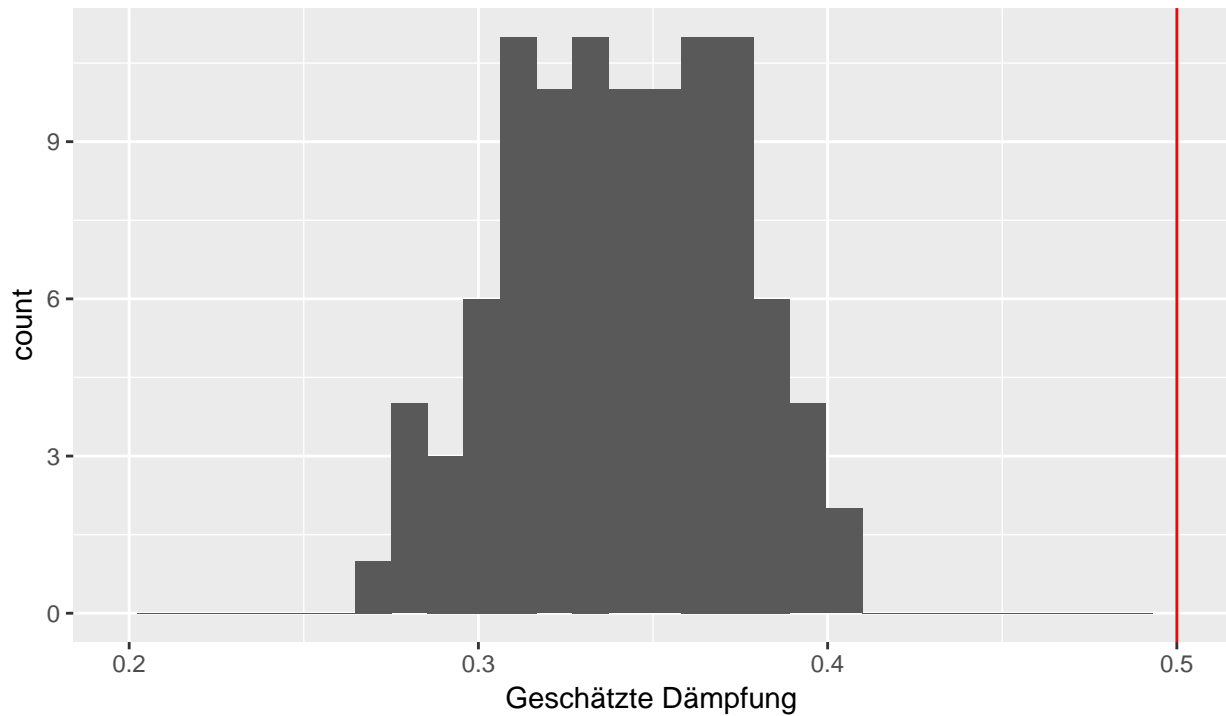
Wahre Dämpfung: 0.7 Auswertungszeitraum: 0 bis 4



Percent of Intervalls (value $\pm 2 \cdot \text{sd}$) that contain the true value: 0

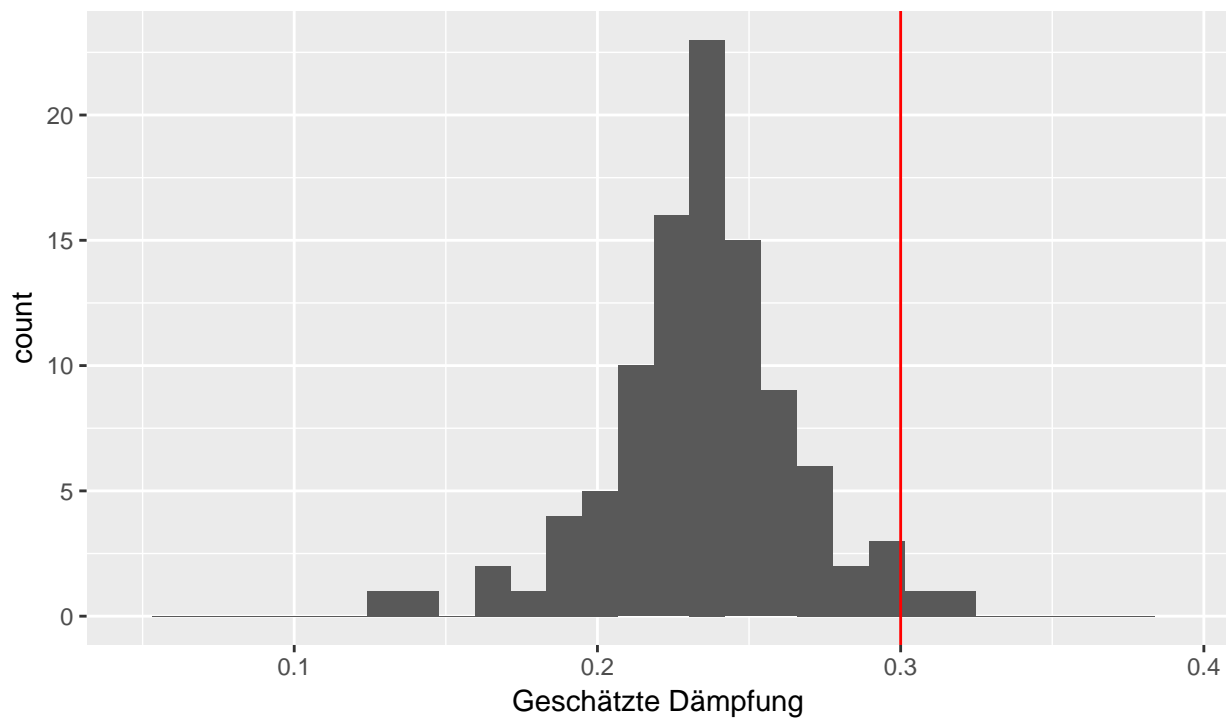
Schätzung von σ mit durch linearen fit an quadriertes log-transformiertes signal

Wahre Dämpfung: 0.5 Auswertungszeitraum: 0 bis 4



Schätzung von σ mit durch linearen fit an quadriertes log-transformiertes signal

Wahre Dämpfung: 0.3 Auswertungszeitraum: 0 bis 4



In allen drei Fällen wird also die Dämpfung systematisch unterschätzt. Der systematische Fehler ist dabei umso größer, je größer die ‘wahre’ Dämpfung ist. Der Grund hierfür ist oben erklärt.

Nichtlinearer Fit von einer gedämpften harmonischen Funktion

In der Annahme das zu untersuchende Signal ist von der Form $ae^{-bt}\sin(2\pi ft)$ versucht man dieses Form an das Signal in den Parametern a, b, f zu fitten.

Im Regelfall konvergiert dieser nichtlineare Fit aber aufgrund der oft vorhandenen Modulation und Signal-Drift im Regelfall nicht. Deswegen ist die kein gangbarer Weg

Aber: Dieser Methode ist die ‘Beste’ falls das Signal tatsächlich eine gedämpfte harmonische Schwingung ist! Falls das Signal aber eine leichte Modulation aufweist, oder die Frequenz wegläuft liefert diese Methode extrem schlechte Schätzungen bzw. konvergiert noch nicht ein mal.

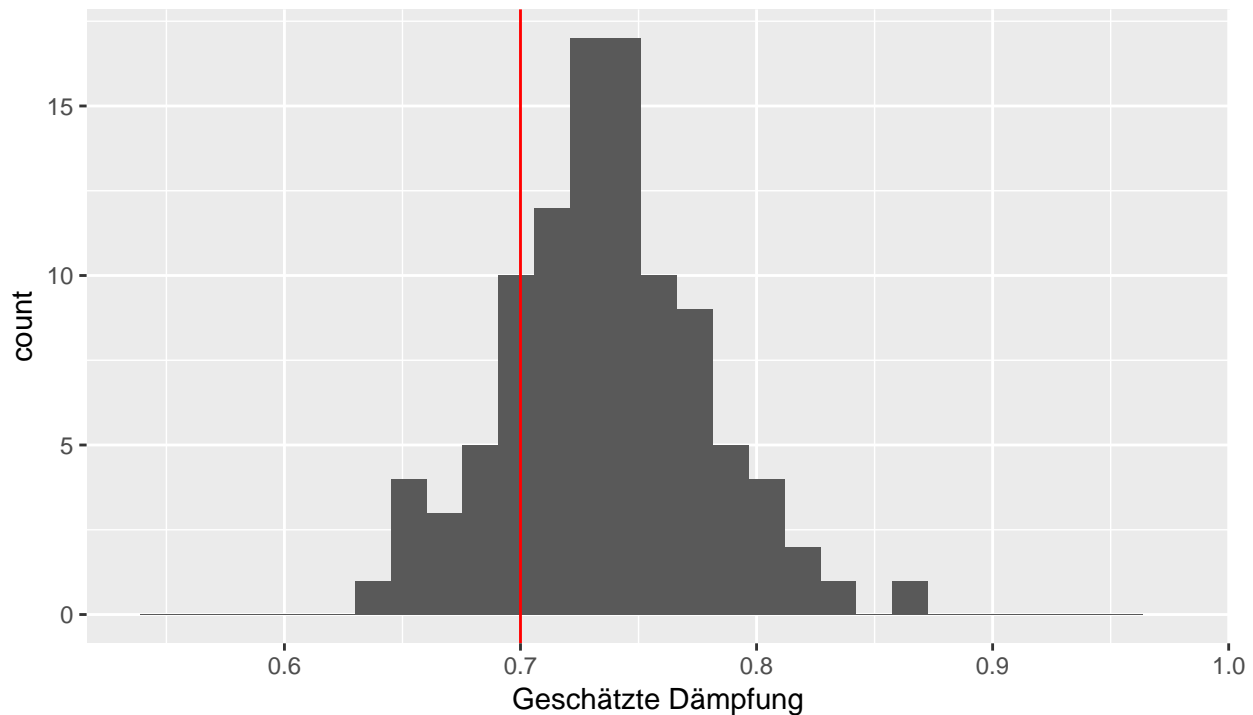
Bestimmung aus der Dämpfung der Amplituden der FT-Koeffizienten

Die Amplituden des Koeffizienten der dominanten Frequenz in aufeinanderfolgenden Zeitfenstern können als Maß für die Signalstärke genommen werden wenn *alle Fenster gleich lang sind (d.h. man hat immer gleichviele Daten)*.

Bemerkung: Offensichtlich macht es Probleme wenn die Anfangsamplituden der gewählten Signale sehr verschieden sind. Dies könnte insbesondere der Fall sein wenn man Radius und die Summe von Radien in die Auswertung nimmt. Eine genauere Untersuchung diese Phänomens steht noch aus.

Schätzung von b durch den gewählten Algorithmus

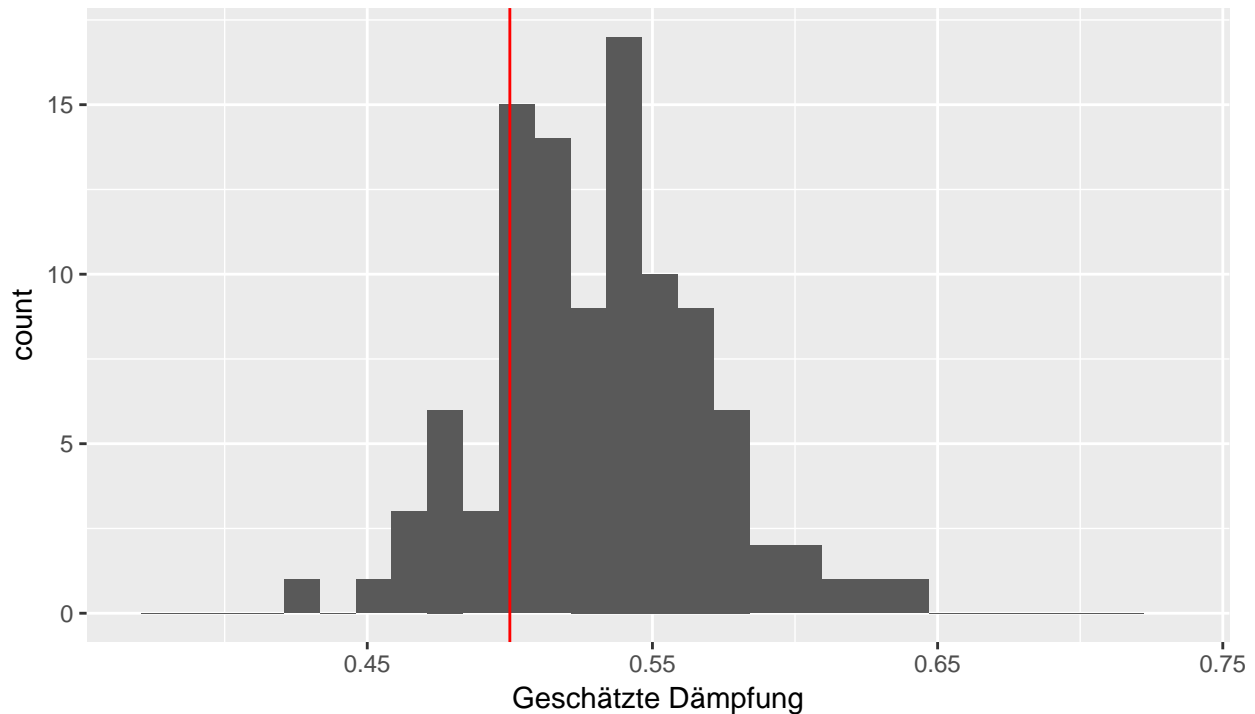
Wahre Dämpfung: 0.7 Auswertungszeitraum: 0 bis 4



Percent of Intervalls (value $\pm 2 \cdot \text{sd}$) that contain the true value: 93.1

Schätzung von b durch den gewählten Algorithmus

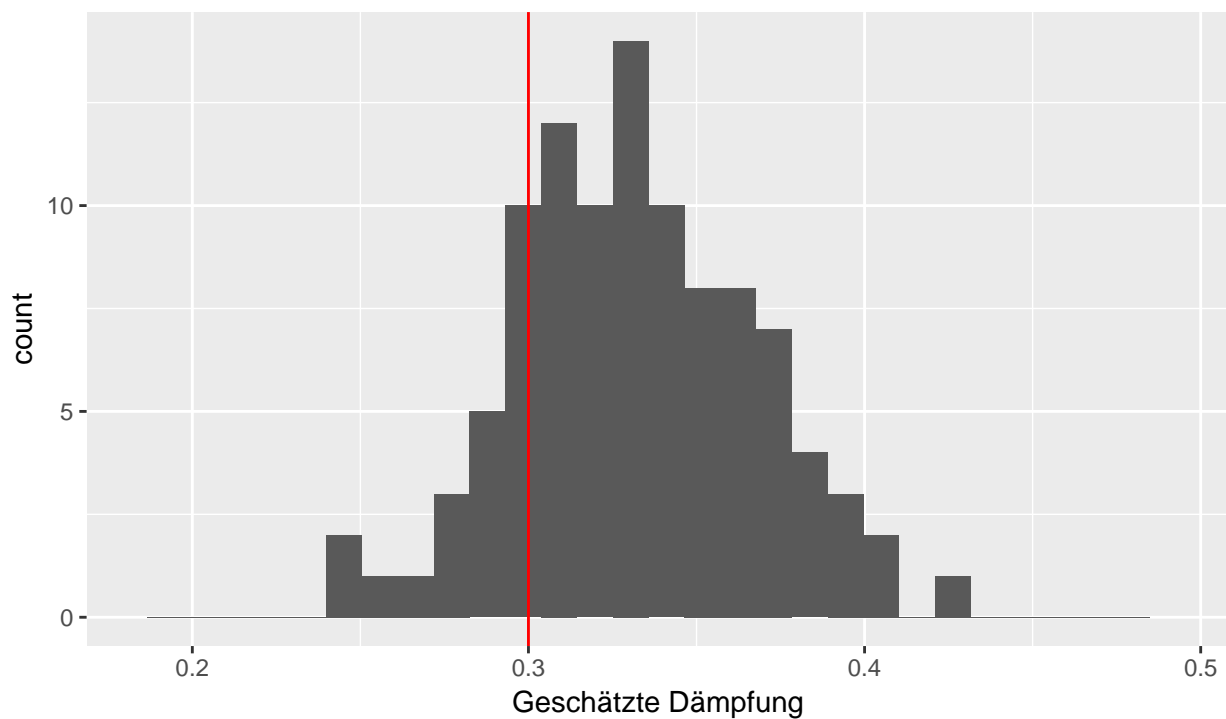
Wahre Dämpfung: 0.5 Auswertungszeitraum: 0 bis 4



Percent of Intervalls (value $\pm 2 \cdot sd$) that contain the true value: 84.2

Schätzung von b durch den gewählten Algorithmus

Wahre Dämpfung: 0.3 Auswertungszeitraum: 0 bis 4



Percent of Intervalls (value $\pm 2 \cdot sd$) that contain the true value: 73.3

Wie man sieht ist auch dieser Algorithmus nicht perfekt, da er die Tendenz hat die Dämpfung zu überschätzen.

D.h. man wird systematisch zu große Viskositäten angeben.

Im Vergleich mit den anderen beiden Methoden ist er aber derjenige mit den besten Eigenschaften. Eine Optimierung dieses Verfahrens um den systematischen Fehler zu verkleinern wäre Aufgabe in der nächsten Programmverbesserung.

Folgerung aus dem Vergleich

Bis auf die direkte Methode haben die Schätzverfahren die Tendenz die Dämpfung systematisch zu unter- oder überschätzen. Insbesondere bei großen Dämpfungen und langen Zeitintervallen wird die Dämpfung unterschätzt, da der deterministische Teil dann auf Null abfällt und das Signal-Rauschen den Fit beeinflusst!

Bei der Wahl der Länge des Zeitfensters zur Auswertung sollte man dies berücksichtigen. Eine automatische Wahl wäre wünschenswert.

- Die Simulationsergebnisse zeigen das die im Programm umgesetzte Art die Dämpfungskonstante zu schätzen nicht so schlecht ist (d.h. besser als die beiden anderen) falls das Signal-rauschverhältnis nicht zu klein wird. D.h. große Dämpfungen bei langen Beobachtungen sind problematisch. Im Programm sind aber heuristische Regeln angegeben wie man kleine Signal/Rausch-Verhältnisse erkennen kann und diese dann aus der Auswertung ausnehmen kann. **Dieser Effekt ist z.B. sehr deutlich im Signal R_RadiusX des Datensatzes "LEK94 ISS Batch-1.2b cycle07 rad-2.dat" zu sehen. Hier mußte die Auswertung bei ca. 12s abgebrochen werden.**
- Die gewählte Methode scheint relativ stabil gegen moderate Modulationen, Frequenzdrifts und multiplikatives und additives Rauschen.
- Bei der Angabe eines Konfidenzintervalles für die geschätzten Dämpfungen und damit für die Viskosität muß man vorsichtig sein, da sich in den Simulationen gezeigt hat das die gewählten 95% Konfidenzintervalle nicht zutreffen. Konservativ sollte man also eher 95% Konfidenzintervall das Intervall Schätzwert $\pm 3 \cdot$ Standarderror angeben.

Offene Fragen

Phase zwischen Ellipsen-Hauptachsen

Wenn eine Ellipse (konstanter Fläche) gefittet wird, **müssen** die Signale 180° phasenverschoben sein! D.h. die Lissajou-Figur sollte eine verdrauschte Ellipse sein! Das ist offensichtlich nicht der Fall!

Wie kann es sein dass ein Radius der Ellipse langsam auf Null schrumpft? Eigentlich dürfte es das nicht geben!

Bei raumfesten Hauptachsen kann es sein das es einen Phasensprung gibt! wenn sich der oszillierende Tropfen wegdreht! Da dann die Zuordnung der Hauptachsen zu den Radien springt!

Ich behaupte das überhaupt keine Ellipse angepasst wird! Es wird der Rand detektiert (pixel-weise! d.h. ohne räumliche Kontinuitäts-Randbedingungen) und dieser wird dann irgendwie geglättet.

Die Signale RadiusX ... sind dann einfach der Abstand vom *geschätzten* Kugel-Schwerpunkt zum Rand.