

Ausbreitung eines Virus

Karl Stroetmann

2. Januar 2006

Wir wollen im folgenden die Ausbreitung eines Virus in einem großen Netzwerk untersuchen. Es bezeichnet x die Prozentzahl der Rechner, die noch nicht mit dem Virus infiziert sind, y ist die Prozentzahl der Rechner, die den Virus verbreiten und z ist die Prozentzahl der Rechner, die abgestürzt sind. Wir machen die Annahme, dass ein Rechner, der infiziert wird, solange andere Rechner kontaktiert und diese mit dem Virus infiziert, bis er auf einen Rechner trifft, der bereits infiziert ist. In diesem Fall stürzen beide Rechner ab. Die Ausbreitung wird dann durch das folgende System von Differentialgleichungen beschrieben:

$$\begin{aligned}\frac{dx}{dt} &= -\lambda * x * y \\ \frac{dy}{dt} &= +\lambda * x * y - 2 * \lambda * y^2 - \lambda * y * z \\ \frac{dz}{dt} &= +2 * \lambda * y^2 + \lambda * y * z\end{aligned}$$

Die Anfangsbedingungen lauten:

$$\begin{aligned}x(0) &= 1 - \varepsilon \\ y(0) &= \varepsilon \\ z(0) &= 0\end{aligned}$$

Hier bezeichnet ε die Prozentzahl der Rechner, die anfangs infiziert sind. Wir definieren $a(t) := x(t) + y(t) + z(t)$. Daraus folgt sofort

$$\frac{da}{dt} = \frac{dx}{dt} + \frac{dy}{dt} + \frac{dz}{dt} = 0.$$

Also ist die Funktion $a(t)$ konstant und es gilt

$$a(t) = a(0) = x(0) + y(0) + z(0) = 1.$$

Damit reicht es aus, das folgende System von DGL zu lösen:

$$\begin{aligned}\frac{dx}{dt} &= -\lambda * x * y \\ \frac{dy}{dt} &= +\lambda * x * y - 2 * \lambda * y^2 - \lambda * y * z\end{aligned}$$

Die Funktion $z(t)$ ergibt sich dann einfach nach der Formel

$$z(t) = a(t) - x(t) - y(t) = 1 - x(t) - y(t).$$

Um dieses System zu lösen, eliminieren wir die Zeit und bilden den Differentialquotienten $\frac{dy}{dx}$:

$$\begin{aligned} \frac{dy}{dx} &= \frac{\frac{dy}{dt}}{\frac{dx}{dt}} \\ &= \frac{+\lambda * x * y - 2 * \lambda * y^2 - \lambda * y * z}{-\lambda * x * y} \\ &= -\frac{x * y - 2 * y^2 - y * (1 - x - y)}{x * y} \\ &= -\frac{x * y - 2 * y^2 - y + y * x + y^2}{x * y} \\ &= -\frac{2 * x * y - y^2 - y}{x * y} \\ &= -\frac{2 * x - y - 1}{x} \end{aligned}$$

Als Anfangsbedingung haben wir dann

$$y(1 - \varepsilon) = \varepsilon.$$

Damit haben wir das ursprüngliche Problem auf eine gewöhnliche DGL reduziert. Um diese DGL weiter zu vereinfachen definieren wir

$$u(x) := y(x) + 1.$$

Dann gilt offenbar

$$\frac{du}{dx} = \frac{dy}{dx}$$

und für u finden wir die DGL

$$\frac{du}{dx} = -2 + \frac{u}{x}.$$

Die Anfangsbedingung für u lautet

$$u(1 - \varepsilon) = y(1 - \varepsilon) + 1 = 1 + \varepsilon.$$

Die DGL für $u(x)$ ist eine inhomogene lineare DGL, allerdings mit einem variablen Koeffizienten. Wir lösen daher zunächst die zugeordnete homogene DGL

$$\frac{du}{dx} = \frac{u}{x}$$

indem wir diese wie folgt umformen:

$$\frac{du}{u} = \frac{dx}{x}.$$

Integration auf beiden Seiten liefert

$$\ln(u) = \ln(x) + c,$$

wobei c eine Integrations-Konstante ist. Wenden wir auf beiden Seiten die Exponential-Funktion an, so erhalten wir

$$u(x) = x * C, \quad \text{mit } C = \exp(c).$$

Gemäß dem Verfahren der Variation der Konstanten machen wir also für $u(x)$ den Ansatz

$$u(x) = x * C(x)$$

und setzen dies in die ursprüngliche DGL für u ein. Dabei erhalten wir

$$C(x) + x * \frac{dC}{dx} = -2 + \frac{x * C(x)}{x}.$$

Dies vereinfacht sich zu

$$x * \frac{dC}{dx} = -2$$

und damit finden wir

$$\frac{dC}{dx} = -\frac{2}{x}.$$

Die Lösung lautet offenbar

$$C(x) = -2 * \ln(x) + \alpha,$$

wobei α eine geeignet zu bestimmende Integrations-Konstante ist. Um α zu bestimmen, setzen wir den Werte für $u(1 - \varepsilon)$ in $u(x) = x * C(x)$ ein und erhalten:

$$1 + \varepsilon = (1 - \varepsilon) * (\alpha - 2 * \ln(1 - \varepsilon)).$$

Das liefert für α sofort

$$\alpha = \frac{1 + \varepsilon}{1 - \varepsilon} + 2 * \ln(1 - \varepsilon).$$

Lassen wir ε gegen 0 gehen, so geht α gegen 1:

$$\lim_{\varepsilon \rightarrow 0} \alpha = 1.$$

Zur Vereinfachung rechnen wir mit diesem Wert weiter. Damit finden wir für $u(x)$

$$\begin{aligned} u(x) &= x * (\alpha - 2 * \ln(x)) \\ &= x * \left(\frac{1 + \varepsilon}{1 - \varepsilon} + 2 * \ln(1 - \varepsilon) - 2 * \ln(x) \right) \\ &\xrightarrow{\varepsilon \rightarrow 0} x * (1 - 2 * \ln(x)) \end{aligned}$$

Wegen $u(x) = y(x) + 1$ haben wir also für $\varepsilon \rightarrow 0$

$$y(x) = -1 + x * (1 - 2 * \ln(x)).$$

Die eigentliche Frage ist, wieviele Rechner letztendlich abstürzen. Mit der Zeit geht die Prozentzahl y gegen 0. Dann ist x die Prozentzahl der Rechner, die nicht infiziert werden und läßt sich also als Lösung der folgenden Gleichung bestimmen:

$$y(x) = 0 \Leftrightarrow 0 = -1 + x * (1 - 2 * \ln(x)).$$

Diese Gleichung lösen wir wie folgt:

$$\begin{aligned} \Leftrightarrow 0 &= -1 + x * (1 - 2 * \ln(x)) \\ \Leftrightarrow 1 &= x * (1 - 2 * \ln(x)) \\ \Leftrightarrow 1 &= e^v * (1 - 2 * v) \quad \text{mit } x = e^v \\ \Leftrightarrow -\frac{1}{2} &= e^v * \left(v - \frac{1}{2} \right) \\ \Leftrightarrow -\frac{1}{2} * e^{-\frac{1}{2}} &= \left(v - \frac{1}{2} \right) * e^{v - \frac{1}{2}} \end{aligned}$$

Die letzte Gleichung hat $v = 0$ als offensichtliche Lösung. Diese Lösung liefert aber nur den Wert $x = 1$, der dem Fall entspricht, dass der Virus sich gar nicht ausbreitet. Um die nicht-triviale Lösung zu finden betrachten wir die in Abbildung 1 gezeigte Funktion $w \mapsto w * e^w$.

Wie sehen, dass es für negative Werte r , die in dem Intervall $[-e^{-1}, 0]$ liegen, zwei Werte von w gibt, für die $w * e^w = r$ ist. Einer dieser Werte ist kleiner als -1 , der andere ist größer. Diese beiden Werte werden durch die beiden *Lambertschen* W-Funktionen W_{-1} und W_0 berechnet. Die für uns interessante W-Funktion ist W_{-1} , denn diese rechnet den kleineren Wert aus, für alle $r \in [-e^{-1}, 0]$ gilt:

$$W_{-1}(r) * e^{W_{-1}(r)} = r \quad \text{und} \quad W_{-1}(r) < e^{-1}.$$

Damit können wir die obige Gleichung wie folgt lösen:

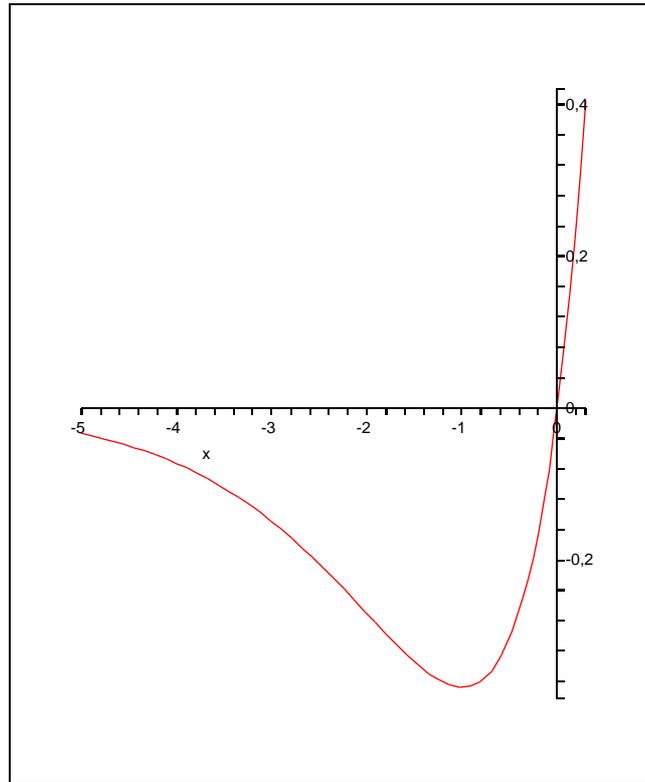


Abbildung 1: Die Funktion $w \mapsto w * e^w$.

$$\begin{aligned}
 -\frac{1}{2} * e^{-\frac{1}{2}} &= (v - \frac{1}{2}) * e^{v - \frac{1}{2}} \\
 \Leftrightarrow W_{-1}(-\frac{1}{2} * e^{-\frac{1}{2}}) &= v - \frac{1}{2} \\
 \Leftrightarrow \frac{1}{2} + W_{-1}(-\frac{1}{2} * e^{-\frac{1}{2}}) &= v \\
 \Leftrightarrow e^{\frac{1}{2} + W_{-1}(-\frac{1}{2} * e^{-\frac{1}{2}})} &= x \quad \text{wegen } x = e^v
 \end{aligned}$$

Die Lambertsche W -Funktion ist in Maple und Mathematica implementiert. Numerisch ergibt sich

$$W_{-1}(-\frac{1}{2} * e^{-\frac{1}{2}}) = -1.756431209.$$

Damit ergibt sich für die Prozentzahl der Rechner, die nicht infiziert werden:

$$x = e^{\frac{1}{2} + W_{-1}(-\frac{1}{2} * e^{-\frac{1}{2}})} = 0.2846681369.$$

Die Prozentzahl, die infiziert werden, ist damit $z = 1 - x = 0.7153318631$.