Freie Universität Berlin
Analytische Mechanik
Sommer-Semester 2020

Übungsblatt 10
Erhaltungsgrößen

Abgabe bis: 03.07.2020 um 12:00 Uhr

---

1. **Erhaltungsgrößen im Zentralpotential** [2+1+2+5+4+1+2=15 Punkte]
In dieser Aufgabe betrachten wir ein Teilchen der Masse m im Zentralpotential $U(\mathbf{r}) = -\frac{k}{|\mathbf{r}|},\ k > 0$.

a) Betrachten Sie zunächste eine infinitesimale Drehung um eine Achse $\mathbf{n}$:

$$\mathbf{r} \to \mathbf{r}^* = \mathbf{r} + (\mathbf{n} \times \mathbf{r})\epsilon, \tag{0.1}$$

und zeigen Sie für eine allgemeine Lagrangefunktion $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}) \to \mathcal{L}^* = \mathcal{L}(\mathbf{r}^*(\epsilon), \dot{\mathbf{r}}^*(\epsilon))$ den folgenden Zusammenhang

$$\left(\frac{\partial \mathcal{L}^*}{\partial \epsilon}\right)_{\epsilon=0} = \frac{d}{dt}\mathbf{n} \cdot (\mathbf{r} \times \mathbf{p}). \tag{0.2}$$

Lösung: Wir haben $\mathcal{L}^* = \mathcal{L}(\mathbf{r} + (\mathbf{n} \times \mathbf{r})\epsilon, \mathbf{r} + (\mathbf{n} \times \dot{\mathbf{r}})\epsilon)$. Damit erhalten wir (mit Summenkonvention)

$$\left(\frac{\partial \mathcal{L}^*(\mathbf{r}^*(\epsilon), \dot{\mathbf{r}}^*(\epsilon))}{\partial \epsilon}\right)_{\epsilon=0} = \frac{\partial \mathcal{L}}{\partial r_i}(\mathbf{n} \times \mathbf{r})_i + \frac{\partial \mathcal{L}}{\partial \dot{r}_i}(\mathbf{n} \times \dot{\mathbf{r}})_i \tag{0.3}$$

$$\overset{\text{ELG}}{=} \left(\frac{d}{dt}p_i\right)(\mathbf{n} \times \mathbf{r})_i + p_i \frac{d}{dt}(\mathbf{n} \times \mathbf{r})_i \tag{0.4}$$

$$= \frac{d}{dt}\mathbf{p} \cdot (\mathbf{n} \times \mathbf{r}) \tag{0.5}$$

$$= \frac{d}{dt}\mathbf{n} \cdot (\mathbf{r} \times \mathbf{p}), \tag{0.6}$$

wobei wir benutzt haben, dass $\frac{\partial \mathcal{L}}{\partial \dot{r}_i} = p_i$.

b) Zeigen Sie mit (0.2), dass im Zentralpotential der Drehimpuls erhalten ist.

Lösung: Betrachten wir zuerst den kinetischen Term, dort haben wir

$$\frac{d}{d\epsilon}T^*\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\frac{m}{2}(\dot{\mathbf{r}}^*)^2 = \frac{d}{d\epsilon}\frac{m}{2}(\dot{\mathbf{r}}^2 + 2\dot{\mathbf{r}} \cdot (\mathbf{n} \times \dot{\mathbf{r}})\epsilon + O(\epsilon^2))\bigg|_{\epsilon=0} \tag{0.7}$$

$$= \frac{d}{d\epsilon}O(\epsilon^2)\bigg|_{\epsilon=0} \tag{0.8}$$

$$= 0. \tag{0.9}$$

Für den Potentialterm gilt zunächst für einen beliebigen Vektor $\mathbf{v}$

$$\frac{d}{d\epsilon}\frac{1}{|\mathbf{r} + \mathbf{v}\epsilon|}\bigg|_{\epsilon=0} = -\frac{\mathbf{v} \cdot \mathbf{r}}{|\mathbf{r}|^3}. \tag{0.10}$$

Damit erhalten wir für den Potentialterm

$$\frac{d}{d\epsilon}U^*\bigg|_{\epsilon=0} = \frac{(\mathbf{n} \times \mathbf{r}) \cdot \mathbf{r}}{|\mathbf{r}|^3} = 0. \tag{0.11}$$

Mit (0.2) gilt nun, dass $\mathbf{n}\cdot(\mathbf{r}\times\mathbf{p})$ für beliebige $\mathbf{n}$ eine Erhaltungsgrösse ist. Insbesondere gilt dies auch für $\mathbf{n} \in \mathcal{B} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$, wo $\mathcal{B}$ eine Basis bildet. Damit muss $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ eine Erhaltungsgröße sein.

c) Zeigen Sie direkt mithilfe der Euler-Lagrange Gleichungen, dass der Drehimpuls eines isotropischen Systems erhalten ist.
Hinweis: Benutzen Sie dazu Zylinderkoordinaten

Lösung: In einem isotropischen System können wir die $z-$Achse in eine beliebige Richtung legen und in Zylinderkoordinaten gehen. In Zylinderkoordinaten (ρ, ϕ, z) haben wir die Lagrangefunktion

$$\mathcal{L} = \frac{m}{2}(\dot{\rho}^2 + r^2\dot{\phi}^2 + \dot{z}^2) + U(r, z), \tag{0.12}$$

wo wegen der Isotropie des Systemes das Potential $U(r, z)$ explizit nicht vom Winkel ϕ abhängt. Mit der ELG für ϕ haben wir dann

$$\frac{d}{dt} mr^2\dot{\phi} = 0. \tag{0.13}$$

Betrachten wir die $z-$Komponente des Drehimpulses in Zylinderkoordinaten

$$\mathbf{r} = \begin{pmatrix} \rho\cos\phi \\ \rho\sin\phi \\ z \end{pmatrix}, \qquad \dot{\mathbf{r}} = \begin{pmatrix} \dot{\rho}\cos\phi - \rho\dot{\phi}\sin\phi \\ \dot{\rho}\sin\phi + \dot{\phi}\cos\phi \\ z \end{pmatrix}, \tag{0.14}$$

$$(\mathbf{r}\times\mathbf{p})_z = m(\mathbf{r}\times\dot{\mathbf{r}})_z = mr^2\dot{\phi}, \tag{0.15}$$

sehen wir, dass diese mit der Erhaltungsgröße aus (0.13) übereinstimmt. Da von vornherein wegen der isotropie des Systemes die Ausrichtung der $z-$Achse beliebig war, ist der Drehimpuls erhalten.

d) Zeigen Sie, dass die Lagrangefunktion des Teilchens im Zentralpotential unter den drei infinitesimalen Koordinatentransformationen ($j = 1,\ 2,\ 3$)

$$x_i \to x_i^*(\epsilon, j) = x_i + \epsilon\left[\dot{x}_i x_j - \frac{1}{2}x_i\dot{x}_j - \frac{1}{2}\delta_{ij}\sum_k x_k\dot{x}_k\right] \tag{0.16}$$

für die kartesischen Komponenten des Ortsvektors ($i = 1,\ 2,\ 3$) bis auf eine totale Zeitableitung invariant ist.
Hinweis: Verwenden Sie die Bewegungsgleichungen.

Lösung: Wir benutzen die Einsteinsche Summenkonvention. Die Lagrangefunktion des Systemes lautet

$$\mathcal{L} = \frac{m}{2}\dot{x}_i^2 + \frac{k}{|\mathbf{x}|}. \tag{0.17}$$

Damit ergeben sich mit ELG die Bewegungsgleichungen

$$m\ddot{x}_i = -\frac{kx_i}{|\mathbf{x}|^3}. \tag{0.18}$$

Die Änderung der Lagrangefunktion in erster Ordnung ergibt sich zu

$$\left.\frac{\partial \mathcal{L}(x^*(\epsilon,j),\dot{x}^*(\epsilon,j))}{\partial\epsilon}\right|_{\epsilon=0} \tag{0.19}$$

$$= \dot{x}_i^*(\epsilon,j)\left.\frac{\partial \dot{x}_i^*(\epsilon,j)}{\partial\epsilon}\right|_{\epsilon=0} - \frac{k x_i^*(\epsilon,j)}{|\mathbf{x}^*(\epsilon,j)|^3}\left.\frac{\partial x_i^*(\epsilon,j)}{\partial\epsilon}\right|_{\epsilon=0} \tag{0.20}$$

$$= m\dot{x}_i\left[\ddot{x}_i x_j + \frac{1}{2}\dot{x}_i\dot{x}_j - \frac{1}{2}x_i\ddot{x}_j - \frac{1}{2}\delta_{ij}(\dot{x}_k)^2 + x_k\ddot{x}_k\right] \tag{0.21}$$

$$\quad - \frac{k x_i}{|\mathbf{x}|^3}\left[\dot{x}_i x_j - \frac{1}{2}x_i\dot{x}_j - \frac{1}{2}\delta_{ij}x_k\dot{x}_k\right] \tag{0.22}$$

$$= \underbrace{(m\ddot{x}_i)}_{-\frac{k}{|\mathbf{x}|^3}x_i}\dot{x}_i x_j + \frac{m}{2}\dot{x}_i^2\dot{x}_j - \frac{1}{2}\underbrace{(m\ddot{x}_j)}_{-\frac{k}{|\mathbf{x}|^3}x_j}x_i\dot{x}_i - \frac{m}{2}\dot{x}_i\delta_{ij}\dot{x}_k^2 + x_k\underbrace{(m\ddot{x}_k)}_{-\frac{k}{|\mathbf{x}|^3}x_k} \tag{0.23}$$

$$\quad - \frac{k}{|\mathbf{x}|^3}x_i\dot{x}_i x_j + \frac{1}{2}\frac{k}{|\mathbf{x}|^3}x_i^2\dot{x}_j + \frac{1}{2}\frac{k}{|\mathbf{x}|^3}x_i\delta_{ij}x_k\dot{x}_k \tag{0.24}$$

$$= -\frac{k}{|\mathbf{x}|^3}(\mathbf{x}\cdot\dot{\mathbf{x}})x_j + \frac{m}{2}\mathbf{x}^2\dot{x}_j + \frac{k}{|\mathbf{x}|^3}(\mathbf{x}\cdot\dot{\mathbf{x}})x_j - \frac{m}{2}\mathbf{x}^2\dot{x}_j + \frac{1}{2}\frac{k}{|\mathbf{x}|^3}\dot{x}_j\mathbf{x}^2 \tag{0.25}$$

$$\quad - \frac{k}{|\mathbf{x}|^3}(\mathbf{x}\cdot\dot{\mathbf{x}})x_j + \frac{1}{2}\frac{k}{|\mathbf{x}|^3}\mathbf{x}^2\dot{x}_j + \frac{1}{2}\frac{k}{|\mathbf{x}|^3}(\mathbf{x}\cdot\dot{\mathbf{x}})x_j \tag{0.26}$$

$$= \frac{k}{|\mathbf{x}|^3}\mathbf{x}^2\dot{x}_j - \frac{k}{|\mathbf{x}|^3}(\mathbf{x}\cdot\dot{\mathbf{x}})x_j \tag{0.27}$$

$$= \frac{k}{|\mathbf{x}|}\dot{x}_j + \left[\frac{d}{dt}\frac{k}{|\mathbf{x}|}\right]x_j \tag{0.28}$$

$$= \frac{d}{dt}\left[\frac{k}{|\mathbf{x}|}x_j\right]. \tag{0.29}$$

Damit sehen wir, dass die Lagrangefunktion bis auf eine totale Zeitableitung $\frac{d}{dt}f_j$ invariant ist, mit

$$f_j(\mathbf{x}) = \frac{k}{|\mathbf{x}|}x_j. \tag{0.30}$$

e) Zeigen Sie, dass die über das Noethertheorem mit den Koordinatentransformationen verknüpften Erhaltungsgrößen die Komponenten des Vektors

$$\mathbf{A} = \dot{\mathbf{x}} \times \mathbf{L} - k\frac{\mathbf{x}}{|\mathbf{x}|} \tag{0.31}$$

sind, wobei $\mathbf{L}$ den Drehimpuls des Teilchens bezeichnet.

Lösung: Laut dem Noethertheorem ist die Erhaltungsgrösse für Invarianz unter $x_i \to x_i^* = x_i + \epsilon\eta_{j,i}$bis auf auf eine totale Zeitableitung $\frac{d}{dt} f_j(\mathbf{x})$ gegeben durch

$$\begin{aligned}
Q_j &= \frac{\partial \mathcal{L}}{\partial \dot{x}_i}\eta_{j,i} - f_j(\mathbf{x}) && (0.32)\\
&= m\dot{\mathbf{x}}^2 x_j - m\dot{x}_j(\mathbf{x}\cdot\dot{\mathbf{x}}) - k\frac{x_j}{|\mathbf{x}|} && (0.33)\\
&= m\left[\mathbf{x}(\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}) - \dot{\mathbf{x}}(\dot{\mathbf{x}}\cdot\mathbf{x})\right]_j - k\frac{x_j}{|\mathbf{x}|} && (0.34)\\
&= m\left[\dot{\mathbf{x}}\times(\mathbf{x}\times\dot{\mathbf{x}})\right]_j - k\frac{x_j}{|\mathbf{x}|} && (0.35)\\
&= \left[\dot{\mathbf{x}}\times\mathbf{L} - k\frac{\mathbf{x}}{|\mathbf{x}|}\right]_j && (0.36)\\
&= A_j. && (0.37)
\end{aligned}$$

In Zeile (0.35) haben wir die $bac{-}cab$ regel rückwärts angewendet. Dies ist der Lenz'sche Vektor, den Sie bereits in einer vergangenen Übung kennengelernt haben.

f) Bestimmen Sie die Hamiltonfunktion des Teilchens im Zentralpotential und zeigen Sie mithilfe der Poisson-Klammer dass die Komponenten des Drehimpulses erhalten sind.

Lösung: Die Hamiltonfunktion eines Teilchens im Zentralpotential bestimmt sich zu

$$H = \frac{\mathbf{p}^2}{2m} - \frac{k}{|\mathbf{q}|}. \tag{0.38}$$

Damit haben wir auch

$$\frac{\partial H}{\partial p_j} = \frac{p_j}{m}, \qquad \frac{\partial H}{\partial q_j} = -\frac{kq_j}{|\mathbf{q}|^3}. \tag{0.39}$$

Für die Komponenten des Drehimpulses $L_i = \epsilon_{ijk}q_j p_k$ gilt (mit Summenkonvention)

$$\frac{\partial L_i}{\partial p_j} = \epsilon_{ikj}q_k, \qquad \frac{\partial L_i}{\partial q_j} = \epsilon_{ijk}p_k. \tag{0.40}$$

Damit wird die Poisson-Klammer zu

$$\begin{aligned}
\{H, L_i\} &= \frac{\partial H}{\partial q_j}\frac{\partial L_i}{\partial p_j} - \frac{\partial L_i}{\partial q_j}\frac{\partial H}{\partial p_j} && (0.41)\\
&= -\frac{k}{|\mathbf{q}|^3}\underbrace{\epsilon_{ikj}q_j q_k}_{=0} - \frac{1}{m}\underbrace{\epsilon_{ijk}p_j p_k}_{=0} && (0.42)\\
&= 0. && (0.43)
\end{aligned}$$

g) Zeigen Sie mithilfe der Jacobi-Identität, dass die Poisson-Klammer $\{A,\, B\}$ zweier Erhaltungsgrößen A, B eines Teilchens eine Erhaltungsgröße ist.

Lösung: Wir betrachten zunächst

$$\begin{aligned}
\frac{\partial}{\partial t}\{A,\, B\} &= \frac{\partial}{\partial t}\left(\frac{\partial A}{\partial q}\frac{\partial B}{\partial p} - \frac{\partial B}{\partial q}\frac{\partial A}{\partial p}\right) && (0.44)\\
&= \left(\frac{\partial^2 A}{\partial q\partial t}\frac{\partial B}{\partial p} + \frac{\partial A}{\partial q}\frac{\partial^2 B}{\partial p\partial t} - \frac{\partial^2 B}{\partial q\partial t}\frac{\partial A}{\partial p} - \frac{\partial B}{\partial q}\frac{\partial^2 A}{\partial p\partial t}\right) && (0.45)\\
&= \left\{\frac{\partial A}{\partial t},\, B\right\} + \left\{A,\, \frac{\partial B}{\partial t}\right\}. && (0.46)
\end{aligned}$$

Damit gilt

$$\frac{d}{dt}\{A,\, B\} = \{\{A,\, B\}\,,\, H\} + \frac{\partial}{\partial t}\{A,\, B\} \tag{0.47}$$

$$= \{A,\, \{B,\, H\}\} - \{B,\, \{A,\, H\}\} - \left\{B,\, \frac{\partial A}{\partial t}\right\} + \left\{A,\, \frac{\partial B}{\partial t}\right\} \tag{0.48}$$

$$= \left\{A,\, \{B,\, H\} + \frac{\partial B}{\partial t}\right\} - \left\{B,\, \{A,\, H\} + \frac{\partial A}{\partial t}\right\} \tag{0.49}$$

$$= \left\{\frac{dA}{dt},\, B\right\} + \left\{A,\, \frac{dB}{dt}\right\}. \tag{0.50}$$

Wir sehen nun, dass die Poisson-Klammer von Erhaltungsgrössen ebenfalls eine Erhaltungsgrösse ist. Zudem kann man erkennen, dass unter dem Produkt $\{\cdot,\, \cdot\}$ die Jacobi-Identität die Produktregel beschreibt.

2. **Legendre Transformation** $[2+1+2+2+2+2 = 11$ Punkte$]$

Die Hamiltonsche Mechanik und die Lagrange Mechanik ergeben sich aus einander, indem die Freiheitsgrade des Systems entweder mit $(q, \dot{q})$ oder (q, p) parametrisiert werden. Dadurch wird die Rolle der Lagrangefunktion $L(q, \dot{q}, t)$ durch die Hamiltonfunktion $H(q, p, t)$ besetzt und umgekehrt.

Die Dualitätsbeziehung zwischen $\dot{q}$ und p und damit L und H beruht auf einem allgemeineren mathematischen Prinzip der Legendre-Dualität (oder noch allgemeiner der Fenchel-Dualität). Diese Dualität ist eines der wichtigsten Konzepte der konvexen Analysis und Theorie der konvexen Optimierung. Neben der Mechanik spielt die Legendre-Dualität auch im Übergang zwischen unterschiedlichen Beschreibungen der Thermodynamik eine wichtige Rolle. Hier wollen wir uns einmal genauer anschauen, was dahinter steckt.

Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ bezeichnet man als *konvex*, wenn für alle $x_1, x_2 \in \mathbb{R}^n$ und $\lambda \in [0, 1]$ gilt, dass $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$. Gilt diese Bedingung als strikte Ungleichung für $\lambda \in (0, 1)$ und $x_1 \neq x_2$ so nennt man f *streng konvex*.

a) Zeigen Sie, dass $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ eine konvexe Funktion ist.

Lösung: Sei $\lambda \in [0, 1]$ und $x_1, x_2 \in \mathbb{R}$. Nach Definition gilt es zu zeigen, dass

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

ist. Für $f(x) = x^2$ haben wir

$$\begin{aligned}
(\lambda x_1 + (1-\lambda)x_2)^2 &\leq \lambda x_1^2 + (1-\lambda)x_2^2 \\
\iff\ 0 &\leq -\lambda^2 x_1^2 - (1-\lambda)^2 x_2^2 - 2\lambda(1-\lambda)x_1 x_2 + \lambda x_1^2 + (1-\lambda)x_2^2 \\
&= \lambda(1-\lambda)x_1^2 + \lambda(1-\lambda)x_2^2 - 2\lambda(1-\lambda)x_1 x_2 \\
&= \lambda(1-\lambda)(x_1^2 + x_2^2 - 2x_1 x_2) \\
&= \lambda(1-\lambda)(x_1 - x_2)^2.
\end{aligned}$$

Dies gilt, da für $\lambda \in [0, 1]$ alle Faktoren in der letzten Zeile positiv sind.

(Zusatzaufgabe: Für multidimensionale quadratische Funktionen gilt allgemeiner: Die Funktion $f : \mathbb{R}^n \to R$, $f(x) = \frac{1}{2} x \cdot Ax$ mit $A \in \mathbb{R}^{n \times n}$ einer symmetrischen, positiv-definiten Matrix ist konvex. Zeigen Sie auch dies.)

Für eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ definieren wir die *Legendre-transformierte* oder *konvex-konjugierte* Funktion

$$f^* : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, \infty\}, \quad x^* \mapsto \sup_{x \in \mathbb{R}^n} (x \cdot x^* - f(x))\,.$$

b) Zeigen Sie, dass auch f^* eine konvexe Funktion ist.

Lösung: Dies gilt tasächlich per Definition bereits für beliebige Funktionen $f : \mathbb{R}^n \to \mathbb{R}$. Wir können die Konvexität direkt überprüfen. Sei $\lambda \in [0,1]$ und $x_1^*, x_2^* \in \mathbb{R}^n$, so gilt

$$\begin{aligned} f^*(\lambda x_1^* + (1-\lambda)x_2^*) &= \sup_{x\in\mathbb{R}^n} (x \cdot (\lambda x_1^* + (1-\lambda)x_2^*) - f(x)) \\ &= \sup_{x\in\mathbb{R}^n} (\lambda\,(x \cdot x_1^* - f(x)) + (1-\lambda)\,(x \cdot x_2^* - f(x))) \\ &\leq \lambda \sup_{x\in\mathbb{R}^n} (x \cdot x_1^* - f(x)) + (1-\lambda) \sup_{x\in\mathbb{R}^n} (x \cdot x_2^* - f(x)) \\ &= \lambda f^*(x_1^*) + (1-\lambda) f^*(x_2^*). \end{aligned}$$

c) Sei f streng konvex und stetig differenzierbar. Argumentieren Sie, dass dann

$$f^*(x^*) = x^* \cdot \tilde{x} - F(\tilde{x}),$$

wobei $\tilde{x}$ durch die Bedingung $x^* = \nabla f(\tilde{x})$ eindeutig festgelegt ist.

Lösung: Sei $g(x) = x^* \cdot x - f(x)$. Die Legendre-Transformierte ist definiert als $f^*(x^*) = \sup_{x\in\mathbb{R}} g(x)$. Ist f stetig differenzierbar, so auch g und es gilt für alle $\tilde{x}$, an dem das Supremum angenommen wird, dass $\nabla g(\tilde{x}) = 0$. Nun ist $\nabla g(\tilde{x}) = \nabla|_{\tilde{x}}(x^* x - f(x)) = x^* - \nabla f(\tilde{x})$ und somit $x^* = \nabla f(\tilde{x})$. Da f streng konvex ist, ist $\nabla f(\tilde{x})$ in allen Komponenten streng monoton steigend, somit ist $\tilde{x}$ durch die Bedingung eindeutig festgelegt und auch sicherlich ein Maximum.

d) Geben Sie eine geometrische Interpretation der Legendre-transformierten Funktion f^* einer streng konvexen, differenzierbaren Funktion $f : \mathbb{R} \to \mathbb{R}$. *Hinweis: Betrachten Sie dazu die Tangenten an dem Graphen der Funktion.*

Lösung: Betrachten wir den Graphen $\{(x, f(x)) \mid x \in \mathbb{R}\}$. Für gegebenes $x^* \in \mathbb{R}$ nimmt die Legendre-Transformierte f^* den Wert des geringsten Abstandes in zwischen dem Graphen von f und der Gerade $x \mapsto x^* x$ mit Steigung x^* an. Wir wir im vorrausgegangenen Aufgabenteil gesehen haben, ist dieser Abstand gerade an dem Punkt $(\tilde{x}, f(\tilde{x}))$ erreicht an dem die Steigung von f gerade x^* ist. Verschieben wir also die Gerade, um $f^*(x^*)$ in y-Richtung, erhalten wir die Tangente $x \mapsto x^* x + f^*(x^*)$ von f an der Stelle $\tilde{x}$. Mit anderen Worten gegeben eine Steigung x^* so gibt f^* den y-Achsenabschnitt der Tangente an f mit dieser Steigung an. Da f streng konvex ist wird jede Steigung nur an genauer einer Stelle angenommen. Der duale Graph $\{(x^*, f(x^a st))\}$ legt dadurch f ebenso eindeutig fest.

e) Zeigen Sie, dass die Legendre-Transformation für stetig differenzierbare, strikt konvexe Funktionen eine Involution ist, d.h. es gilt $(f^*)^*(x) = f(x)$.

Lösung: Wir wissen, dass $f^*(x^*) = x^*\tilde{x} - f(\tilde{x})$ mit $x^* = \nabla f(\tilde{x})$ ist. Wenden wir nun wiederum die Legendre-Transformation an, so erhalten wir

$$(f^*)^*(x) = x \cdot \tilde{x}^* - \tilde{x}^* \cdot \tilde{x} + f(\tilde{x})$$

mit $x = \nabla_{x^*}|_{\tilde{x}^*}(x^* \cdot \tilde{x} - f(\tilde{x})) = \tilde{x}$. Einsetzen der Bedingung $\tilde{x} = x$ ergibt, dass $(f^*)^*(x) = f(x)$ ist.

f) Berechnen Sie die Legendre-Transformierte von $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^4/4 - x^2/2$ und $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3/3 + x^2/2$.

Lösung: Wir wollen die Legendre-Transformierte der Funktion $f(x) = x^4/4 - x^2/2$ berechnen. Diese ist stetig differenzierbar, jedoch nicht konvex. Als erstes können wir ausnutzen, dass $f(x) = f(-x)$ ist. Somit ist auch $f^*(-x^*) = \sup_{x\in\mathbb{R}}\{-xx^* - f(x)\} = \sup_{-x\in\mathbb{R}}\{xx^* - f(-x)\} = \sup_{x\in\mathbb{R}}\{xx^* - f(x)\} = f^*(x^*)$. Es genügt folglich f^* für positive x^* zu bestimmen.

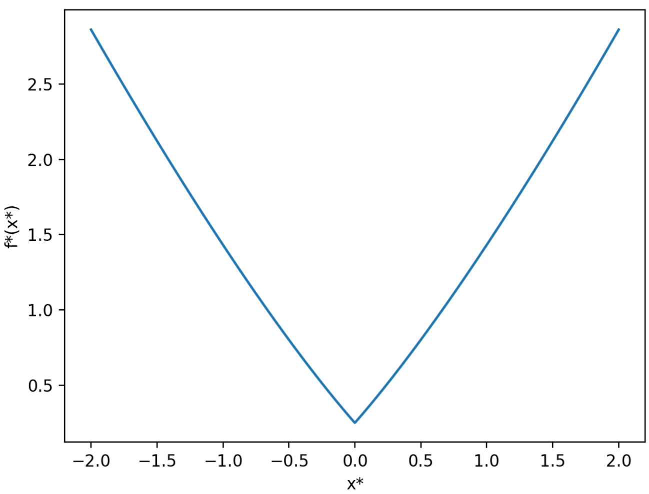
Sei $g_{x^*}(x) = xx^* - f(x)$. Für $x^* = 0$ bestimmt man schnell, dass $g_0(x)$ zwei Maxima bei ± 1 mit Wert $1/4$ hat. Für $x^* > 0$ wird das Maximum bei 1 angehoben und zu größeren Werten von x verschoben. Das Maximum bei negativen Werten von x wird kleiner. Es genügt also den Wert des Maximums bei positiven Werten von x zu bestimmen. Dazu lösen wir $g'_{x^*}(x) = x^* - x^3 + x = 0$ für $x > 0$ und finden den etwas unhandlichen Ausdruck

$$x_{\mathsf{max}} = \frac{2\sqrt[3]{3} + \sqrt[3]{2t^2}}{\sqrt[3]{36\,t}} \tag{0.51}$$

mit $t = 9x^* + \sqrt{-12 + (9x^*)^2}$. Man kann überprüfen, dass für $x^* \geq 0$ dieser Ausdruck in der Tat reell bleibt und eingesetzt in $g''_{x^*}(x)$ ein Maximum beschreibt. Wir versichern uns weiterhin, dass wie erwartet auch $x_{\mathsf{max}} = 1$ für $x^* = 0$ gilt. Somit ist die Legendre-Transformierte geben als

$$f^*(x^*) = g_{x^*}(x_{\mathsf{max}}). \tag{0.52}$$

Da der resultierende Ausdruck ist entsprechend kompliziert und wir begnügen uns mit einem Plot:



Die Legendre-Transformation von $f(x) = x^3/3 + x^2/2$ ist deutlich schneller bestimmt. Für alle x^* gilt, dass $\lim_{x\to\infty}(x^*x - x^3/3 + x^2/2) = \infty$. Das Supremum in der Definition der Legendre-Transformation ist also immer unbeschränkt und $f^*(x^*) = \infty$ für alle x^*.