PLANCKS

Physics League Across Numerous Countries for Kick-Ass Students

Munich & online

7th May 2022

Problems & Solutions

Preliminary remarks

Dear participants,

welcome to this year's PLANCKS exercise booklet! Today, it's all about solving (theoretical) physics problems. The highest scoring teams will hold the title of this year's PLANCKS champions. Before solving the problems, please carefully read the following remarks and excerpts from the PLANCKS rules containing important information about the exam:

1. The competition exam lasts **four hours**. There are **ten problems** each worth **ten points**.
2. The exercises are solved independently and without any external help by each team.
3. The teams commit to sticking to the rules, especially to fairness towards other teams in the scientific contest. If a team violates the rules, it will be disqualified.
4. In case the formulation of an exercise is unclear, every participant can request clarifications from the jury **in written form** by handing the question to the volunteers. The jury then answers the question in written form. If the information is relevant for all teams, the jury will inform every team.
5. The exercises are formulated in English. The solutions need to be handed in **in English** as well.
6. The jury has the right to change or to withdraw problems during the competition. In such a case, the jury informs every team and adjusts the grading appropriately. There are no further consequences.
7. Usage of hardware which is not approved by the jury is forbidden. Dictionaries and non-graphical calculators are allowed!
8. The organisers make sure in the best way possible that the participating teams have no access to mobile devices during the competition.
9. In exceptional cases, the Organising Committee has the right to stop, to interrupt or to extend the competition or to change the grading.
10. In case of dubiety concerning the rules and standards, the jury decides about those.

Last, please note the following:

- Solutions for different exercises have to be handed in on **separate pieces of paper**. This means that you should start a new sheet when writing down the solution of a new exercise.
- Unreadable solutions won't be corrected. If there are two solutions for one exercise and no one is crossed out, both will be counted as wrong.
- Please write the name of your team on **every** sheet of paper you hand in.

If any questions arise during the competition, please ask the assistants on your floor.

We wish you a nice time and lots of success solving the PLANCKS 2022 problems!

The PLANCKS 2022 Jury (Oliver Diekmann, Max Fahn, Miriam Gerharz, Philipp Heinen, Sören Kotlewski, Dr. Charlotta Lorenz, Alexander Osterkorn, Dominik Rattenbacher, Dr. Markus Schmitt, Philippe Suchsland, Rajshree Swarnkar, Dr. Matthias Zimmermann)

Contents

Problem 1

On the Gravitational Three-Body Problem

Prof. Dr. Karl-Henning Rehren – *Georg-August-Universität Göttingen*

Background One sometimes reads that in gravitating system of N stars, each star "orbits around the common center of mass". This statement is certainly to a large extent true if "around" is meant in a qualitative sense. But of course, the orbit is neither a circle nor a Keplerian ellipse with the center of mass in its focal point, so that the notion "around" cannot be given a sharp meaning. On the other hand, for three stars it is known that there exist very special solutions with the stars in an equilateral triangle position that rigidly rotates around its center of mass (each star orbiting in a circle) – independent of the masses.

We want to study (among other things) whether there is in general a single point $\vec{x}^*$ in space such that the total force vectors

$$\vec{F}_i = \sum_{j \neq i} \vec{F}_{i,j} \tag{1.1}$$

of each star point towards that point, and if so, whether this is a special feature of the $1/r^2$-law of the gravitational force. For the sake of this problem, we put Newton's constant $G_N = 1$ and write the force exerted by star 2 on star 1 as

$$\vec{F}_{1,2} = -m_1 m_2 \frac{\vec{x}_{12}}{|\vec{x}_{12}|^{n+1}} \qquad (\vec{x}_{ij} := \vec{x}_i - \vec{x}_j). \tag{1.2}$$

To begin with, we put $n = 2$ (Newton's Law). It is not difficult to show that for $N \geq 4$ stars, a point $\vec{x}^*$ as specified above does not exist in general. (One may just think of systems subdivided into two distant clusters.) So, we consider only $N = 3$.

a) [*1 point*] Recall that the center of mass can be defined as the unique point $\vec{x}_0$ such that

$$\sum_{i=1}^{3} m_i \cdot (\vec{x}_i - \vec{x}_0) = 0. \tag{1.3}$$

For the problem at hand, it is instead useful to define the "center of pseudo-mass" as the unique point $\vec{x}^*$ such that

$$\sum_{i=1}^{3} m_i^* \cdot (\vec{x}_i - \vec{x}^*) = 0, \tag{1.4}$$

where the (configuration-dependent) "pseudo-masses" are defined as

$$m_i^* := |\vec{x}_{jk}|^3 \cdot m_i \tag{1.5}$$

where $j, k = 2, 3$ if $i = 1$ and so on. Solve the equations (1.3) and (1.4) for $\vec{x}_0$ and $\vec{x}^*$, respectively. Can one expect that $\vec{x}^* = \vec{x}_0$ in general?

b) [*3 points*] Show that all total forces $\vec{F}_i$ point towards the center of pseudo-mass $\vec{x}^*$.

Hint: *Use the result of part a) to show that* $\vec{x}_i - \vec{x}^*$ *is a multiple of* $\vec{F}_i$. *It is sufficient to do this for* $i = 1$.

c) [*4 points*] Until this point, we have looked at a fixed instant of time. The next question is whether the situation is stable in time, i.e., whether the triangle formed by the three stars can rigidly rotate around the center of pseudo-mass. By the centrifugal law, this would require that

$$\vec{F}_i = -m_i \omega^2 \cdot \left(\vec{x}_i - \vec{x}^*\right) \tag{1.6}$$

with a common ω^2 for all $i = 1, 2, 3$. Analyze this condition (using the result of b))!

d) [*1 point*] What changes in parts a) to c) if Newton's $1/r^2$-law (1.2) were replaced by a $1/r^n$-law $(n \neq 0)$, i.e. $|\vec{x}_{ij}|^3$ is replaced by $|\vec{x}_{ij}|^{n+1}$ in (1.2) and in the definition of the pseudo-mass?

e) [*1 point*] What changes in parts a) to c) if $n = -1$, i.e., Newton's law (1.2) becomes the elastic force $\vec{F}_{1,2} = -m_1 m_2 \vec{x}_{1,2}$ proportional to the product of masses?

Solution

a) Rearranging (1.3), one gets

$$\vec{x}_0 = \frac{\sum_i m_i \vec{x}_i}{\sum_i m_i}. \tag{1.7}$$

Similarly,

$$\vec{x}^* = \frac{\sum_i m_i^* \vec{x}_i}{\sum_i m_i^*}. \tag{1.8}$$

Because the positions $\vec{x}_i$ are averaged within general different relative weights (unless the three distances are the same), the result will in general be different.

b) By a) and straight-forward computation, we have

$$\begin{aligned}
\vec{x}_1 - \vec{x}^* &= \vec{x}_1 - \frac{m_1 x_{23}^3 \vec{x}_1 + m_2 x_{13}^3 \vec{x}_2 + m_3 x_{12}^3 \vec{x}_3}{m_1 x_{23}^3 + m_2 x_{13}^3 + m_3 x_{12}^3} \\
&= \frac{1}{M^*}\Big(\big(m_1 x_{23}^3 + m_2 x_{13}^3 + m_3 x_{12}^3\big)\vec{x}_1 - m_1 x_{23}^3 \vec{x}_1 - m_2 x_{13}^3 \vec{x}_2 - m_3 x_{12}^3 \vec{x}_3 \Big) \\
&= \frac{1}{M^*}\big(x_{13}^3 m_2 \vec{x}_{12} + x_{12}^3 m_3 \vec{x}_{13} \big) \\
&= -\frac{x_{12}^3 x_{13}^3}{M^* m_1} \cdot \vec{F}_1,
\end{aligned} \tag{1.9}$$

where $M^* = m_1 x_{23}^3 + m_2 x_{13}^3 + m_3 x_{12}^3$ is an abbreviation for the total pseudo-mass. Thus, the force is anti-parallel to the distance from $\vec{x}^*$. Same for $i = 2, 3$.

c) From b) and (1.6) we conclude:

$$\frac{x_{12}^3 x_{13}^3}{m_1 M^*} \overset{!}{=} \frac{1}{m_1 \omega^2} \qquad \Leftrightarrow \qquad \omega^2 = \frac{M^*}{x_{12}^3 x_{13}^3} = \mu \cdot x_{23}^3 \qquad \left(\mu := \frac{M^*}{x_{12}^3 x_{13}^3 x_{23}^3} \right). \tag{1.10}$$

Similar for $i = 2, 3$. Because μ is symmetric under permutations, and ω must be the same for all i, we conclude $|\vec{x}_{12}|^3 = |\vec{x}_{13}|^3 = |\vec{x}_{23}|^3$, i.e., the stars must form an <u>equilateral</u> triangle. In this case, the center of pseudo-mass equals the center of mass. This is the known solution. There is no other one. The permutation-symmetric quantity μ and the side length of the triangle determine the period.

d) Nothing changes. Mutatis mutandis the same results remain true: All forces point towards the center of pseudo-mass, which is in general different from the center of mass. A rigid rotation is possible only for equilateral triangles with arbitrary masses.

e) In part a): Because 3 is replaced by $n + 1 = 0$, the pseudo-mass equals the true mass, and the center of pseudo-mass coincides with the center of mass.
In part b): Nothing changes. The forces point to the center of mass. μ is the total mass.
In part c): The condition on the distance becomes $x_{12}^0 = x_{13}^0 = x_{23}^0$, i.e., $1 = 1 = 1$. This is true <u>independent</u> of the distances. Thus <u>every</u> triangle can rigidly rotate around its center of mass.

Problem 2

James Bond's Car Crash in *Casino Royale*

Dr. Charlotta Lorenz, Dr. Sophie-Charlotte August, Prof. Dr. Sarah Köster – *U Göttingen*

Background Prof. Dr. Metin Tolan, physicist and president of the University of Göttingen, analyzes in his book *Shaken, not stirred – James Bond in the spotlight of physics* a scene from the James Bond movie *Casino Royale*: James Bond is chasing the villain Le Chiffre in a car. Suddenly, Vesper Lynd, Bond's girlfriend, is tied up on the road and Bond jerks the steering wheel to the left, whereupon the car overturns. The following figure shows single frames from the scene in which the car overturns. In the following task, we want to analyze whether the car can actually roll over under the given circumstances.

The car has a mass $m = 1750$ kg, is $b = 1.90$ m wide and $h = 1.28$ m high. It drives with a speed $v = 128$ km/h in the gravitational field of the earth with $g = 9.81$ m/s^2.

a) [*1 point*] Model the car as a cuboid with a center of mass located centrally at the lower third of the car. Sketch the cross-section of the car with all the forces acting on it when the steering wheel is jerked around. Draw the point about which the car rotates when it rolls over sideways. We are only considering rotation along the longitudinal axis of the car in this task.

b) [*1 point*] Calculate the forces and torques acting on the car.

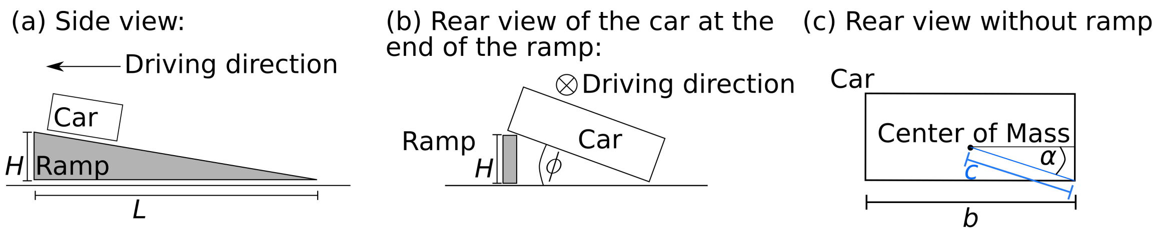
c) [*0.5 points*] Think about what has to happen for the car to start rolling over instead of just driving a turn.

d) [*1 point*] What is the minimum radius of the curve that the car must have before it starts to roll over?

e) [*0.5 points*] In the film, the curve radius is about 200 m, so the car should not roll over by itself. What must happen at a constant curve radius to allow the car to roll over anyway?

f) [*0.5 points*] When the scene was shot, two additional measures were taken to make the car roll over. One additional device was a $L = 2.5$ m long ramp, rising from 0 cm to $H = 10$ cm, over which the left side of the car drove (that is, the car goes up the ramp with its left wheels). The following figure is a simplified representation of the situation. Sketch the car on the ramp from the rear view, and draw the centrifugal force and gravity with their point of application.

g) [*1 point*] Calculate the change in angular velocity ($\dot{\phi}$, see figure (b), where ϕ is drawn) of the car caused by the ramp. Consider how long the car needs to drive over the ramp. For simplicity, assume



Parameters in this sketch are not to scale!
The ramp and the angle ϕ are drawn much larger for clarification than they actually are.

that the height of the ramp can be neglected compared to its length. Then calculate the change in angle of the left side of the car compared to the right side. Use small angle approximation.

h) [*0.5 points*] What does the change in angular velocity mean for angular momentum and torque of the car? Describe the torque as a function of the car's moment of inertia I and angular acceleration.

i) [*1 point*] Now, to calculate the additional force acting on the left side of the car through the ramp, we must first calculate the moment of inertia of the car with respect to the right tires around which it rotates. Model the car as a cuboid. You do not need to explicitly derive the moment of inertia for a cuboid; you can use the familiar formula. Use Steiner's theorem.

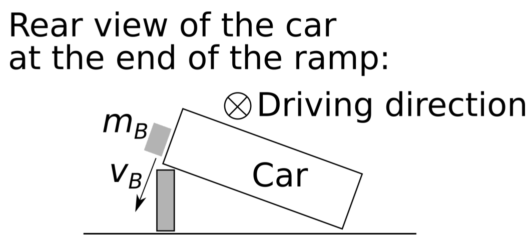
j) [*1 point*] The torque engages the left side of the vehicle. Show that the factor

$$b/(c\cos\alpha)$$

is needed to convert the force from the left side of the vehicle to the center of gravity. b, c and α are shown in the sketch above in (c). Use the lever laws and write the cross product as $\vec{d}\times\vec{f} = |\vec{d}||\vec{f}|\sin\alpha$ with the angle α between the vectors. You can neglect ϕ here compared to α.

k) [*1 point*] Explicitly calculate the additional force F_R generated by the ramp on the center of gravity. Since the ramp is not very high, we can again assume $\alpha >> \phi$ and $\alpha = 24°$.

l) [*1 point*] This force is not sufficient to make the car turn. The film producers tried it and for some cars it works, but for the car used here it is not sufficient. Thus, as a second measure to make the car overturn, an iron bolt with a mass $m_B = 20$ kg is accelerated by $\Delta v_B = 10$ m/s within $\Delta t_B = 0.1$ s and launched near the left wheels of the car as sketched in the following figure. Calculate the additional force on the center of mass generated in this way. Note that here you must again apply the laws of leverage as calculated in (j).

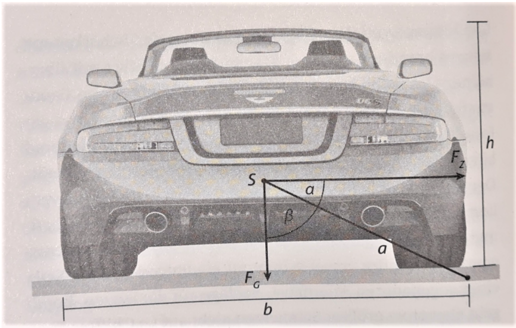


The ramp together with the iron bolt could finally make the car overturn.

Solution

YouTube link: `https://www.youtube.com/watch?v=x-21uPJGXFQ`

a) 0.5 points for correct drawing of the forces; 0.5 points for the point of application of the forces. $a = c$, b and h do not have to be drawn in.



b) Forces: centrifugal force $F_Z = mv^2/r$, gravity force $F_G = mg$. The torques of the respective forces act on the center of gravity of the car. We define $c = \sqrt{(h/3)^2 + (b/2)^2}$. (0.5 points for correct forces; 0.5 points for correct torques)

$$M_Z = F_Z c \sin\alpha\,, \tag{2.1}$$
$$M_G = F_G c \sin\beta\,. \tag{2.2}$$

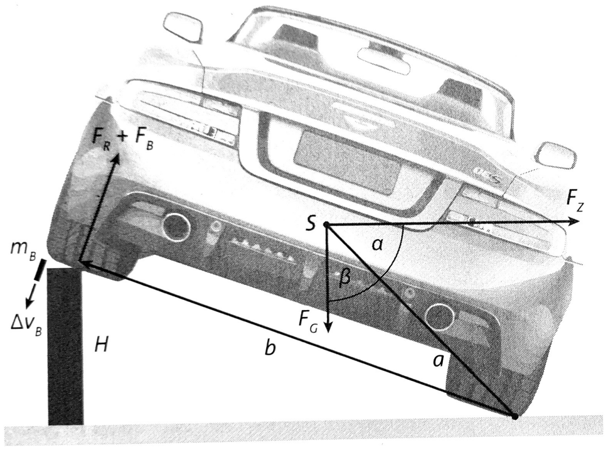
c) The torque pulling the car outward must be greater than the torque pulling the car downward, so $M_Z > M_G$. (0.5 points)

d) After reasoning in (c) and substituting in (b) and transforming, we get (0.5 points for correct calculation; 0.5 point for correct result):

$$r < v^2 \tan\alpha/g \approx 57\,\mathrm{m}\,. \tag{2.3}$$

e) Important key points in the solution: larger torque, change of angular momentum or similar. (0.5 points for correct idea)

f) Sketch (here the complete sketch from the book is shown. 0.5 points for sketching correct forces and correct center of mass):



g) Time it takes the car to go over the ramp (0.5 points for calculations and correct idea; 0.5 points for correct result):

$$\delta t = \frac{L}{v}\,.$$

Angle change:

$$\sin\Delta\phi = \frac{H}{b} \approx \Delta\phi\,.$$

Thus the change in angular velocity:

$$\Delta\omega = \frac{\Delta\phi}{\Delta t} = \frac{vH}{Lb}$$

h) Angular momentum changes so that torque > 0 is created. The torque D is created (0.5 points):

$$D = I\frac{\delta\omega}{\delta t}$$

i) Moment of inertia with $c = \sqrt{(h/3)^2 + (b/2)^2}$ (0.5 points for correct ansatz; 0.5 points for correct solution):

$$I = \frac{1}{12}m(b^2 + h^2) + mc^2\,.$$

j) According to the lever laws, the sum of the total torques acting must be 0, or the torque generated by the force $\vec{F}_L$ on the left side of the vehicle must be equal to the torque generated by the force $\vec{F}_R$ on the center of gravity. That is, if the angle ϕ is negligible with respect to α (0.5 points for correct ansatz, 0.5 points for correct result):

$$\begin{aligned}
\vec{F}_L \times \vec{b} &= -\vec{F}_R \times \vec{c} \\
F_L b &= -F_R c \sin(90^\circ - \alpha) \\
F_L b &= F_R c \cos\alpha \\
\Rightarrow \frac{F_R}{F_L} &= \frac{b}{c\cos\alpha}\,.
\end{aligned}$$

k) (0.5 points for correct ansatz, 0.5 points for correct result)

$$F_R = \frac{D}{r} = \frac{IHv^2}{cbL^2 \cos\alpha} \approx 11600 \text{ N} \ .$$

l) With Newton you get (0.5 points):

$$F_{B,b} = m_B \frac{\delta v_B}{\delta t_B} = 2000 \text{ N} \, .$$

In terms of the center of gravity, this corresponds to (0.5 points):

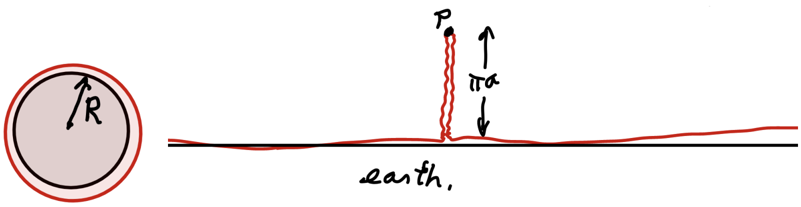
$$F_B = F_{B,b} \frac{b}{c \cos\alpha} \approx 4000 \text{ N}.$$

Problem 3

Rope around the World

Prof. Dr. David DiVincenzo[1], Philippe Suchsland[2] – [1]*Peter Grünberg Institute (PGI-2), Forschungszentrum Jülich, Jülich, Germany,* [2]*Max Planck Institute for the Physics of Complex systems, Dresden, Germany*

Background A loop of rope goes around the circumference of the earth with radius R. The rope is ideal (∞ stretching modulus, ∞ strength, 0 bending modulus, infinitely thin). Its mass density is ρ (kg/m). Its length is such it can be held distance a above the surface, all the way around. This means that if it is laid straight on the surface, there is a $2\pi a$ leftover. We assume $a \ll R$ (e.g. $a = 1\,\mathrm{m}$, $R \sim 10^6\,\mathrm{m}$).



Someone grasps the rope at point P and raises it. What is the maximum height h to which it can be raised?

a) [*1 point*] In a first step, we want to build up some physical intuition. First, introduce the quantity l, the length of the rope not resting on earth when the rope is held at height h. Justify by dimensional analysis and by analysing the limiting behaviour for $a/R \to 0$ in the case $a = \mathrm{const}$ or $\mathrm{R} = \mathrm{const}$ that h, l can be expressed by

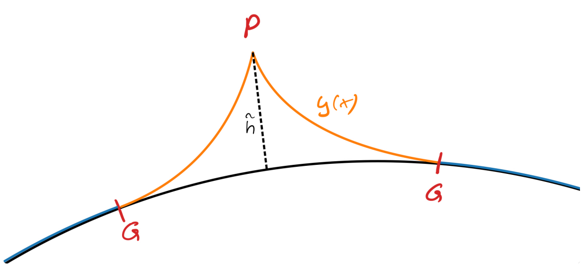
$$l = c_1 a^{1-\alpha} R^{\alpha} + \ldots, \quad 0 < \alpha < 1, \tag{3.1}$$
$$h = c_2 a^{1-\beta} R^{\beta} + \ldots, \quad 0 < \beta < 1, \tag{3.2}$$

in the limit $a/R \ll 1$.

b) [*0.5 points*] Show that $l/R \ll 1$ for $a/R \ll 1$.

c) [*2.5 points*] Calculate the constants c_1, c_2, α, β by calculating l, h in the limit $a/R \ll 1$.
Hint: *Except for $l/R \ll 1$ this task does not rely on the previous ones.*

We now calculate tensions. Suppose the rope is raised to height $\tilde{h}$, $\pi a < \tilde{h} \ll h$. The rope will look something like this (not to scale):



Cable theory shows that $y(x)$, a so-called "catenary curve", is a segment of hyperbolic-cosine curves in the limit of uniform gravitational field. We aim to find the differential equations determining the curves $y(x)$. For that, consider the forces acting on the rope. We work in the limit of $a/R \to \infty$ so you may ignore the curvature of the earth and variations in the gravitational force in all following subtasks.

d) [*2 points*] Derive, but do not solve, a set of three differential equations determining the three unknown functions $y(x)$, $F_{\mathrm{R},x}$ and $F_{\mathrm{R},y}$, where $F_{\mathrm{R},x}$ and $F_{\mathrm{R},y}$ are the two vector components of the tension of the rope $\vec{F}_{\mathrm{R}}$. You are free to choose the reference system and express the strength and direction of the gravitational force as $\vec{g}$.

Hint: *Consider the forces acting on a short section of the rope.*

e) [*1.5 points*] By introducing the points G, where the rope begins to rest on the ground, write down the boundary conditions for the differential equations derived in the previous task, i.e., the conditions needed to fully determine the slope. Particularly, what is the force law that determines the condition at the grounding points G?

f) [*1 point*] Solve the set of differential equations.

Hint: *You might find the solution of the integral*

$$\int \mathrm{d}z \frac{1}{\sqrt{1+z^2}} = \operatorname{asinh}(z) + c, \tag{3.3}$$

where $\operatorname{asinh}(z)$ *is the inverse function of* $\sinh(z)$*, useful.*

The solution of the differential equation takes the form $y(x) = a + \cosh(b(x-c))/b$, which you can use in the following.

g) [*1 point*] We raise P much higher, but still much lower than the maximum; in particular $\tilde{h} = h/10$. What is the tension of the rope at point P?

Hint: *You may assume* $|y'| \ll 1$ *at point* P.

h) [*0.5 points*] Based on the result of the previous exercise and h from part c), what will happen with the rope if we raise a real rope in the limit $R \to \infty$ to height $\tilde{h} = h/10$?

Solution

a) As a first step as given in the exercise, we assume that h, l are analytical and, hence, can be expressed as

$$l = c_1 R(a/R)^{1-\alpha} + \mathcal{O}((a/R)^{1-\alpha+\epsilon_1}) \tag{3.4}$$

$$h = c_2 R(a/R)^{1-\beta} + \mathcal{O}((a/R)^{1-\beta+\epsilon_2}) \tag{3.5}$$

for $a/R \ll 1$ with some $\alpha, \beta, \epsilon_1, \epsilon_2 > 0$ and dimensionless constants c_1, c_2.

The limit $a/R \to 0$ can be approached by $a = \text{const}, R \to \infty$ or $a \to 0, R = \text{const}$. In the limit of $a \to 0$ the rope fits exactly one time around the earth and, hence, we expect $h \to 0, l \to 0$, as h, l are continues functions and at $a = 0$ $h = 0, l = 0$. In this limit for $R = \text{const}$ the functions have to obey

$$l \sim a^{\alpha'}, h \sim a^{\beta'} \tag{3.6}$$

with $\alpha', \beta' > 0$ in order to approach 0 for $a \to 0$.

In the limit $a = \text{const}, R \to \infty$ we expect both, h and l, to diverge towards infinity as well. In order to see that we provide a lower bound on h which diverges. For that consider the construction as shown in Fig. 3.1. Clearly, the maximum h is larger than the $\tilde{h}$. In the limit $R \to \infty$, higher orders in

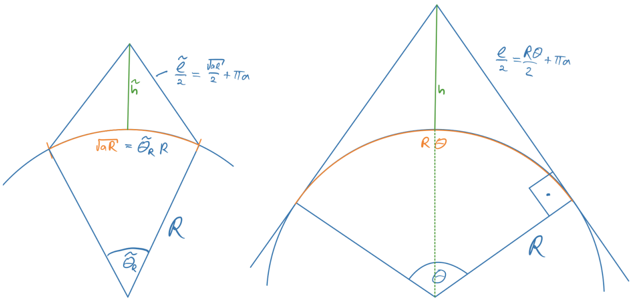


Figure 3.1 Construction for the lower bound on h in the limit $R \to \infty, a = \text{const}$.

$\tilde{\theta}_R$ vanish and we can write

$$\tilde{h} = \sqrt{(\sqrt{aR}/2 + \pi a)^2 - aR/4} = \sqrt{\pi a \sqrt{aR} + \pi^2 a^2}. \tag{3.7}$$

Hence, $\tilde{h} \to \infty$ and therefore $h \to \infty$ and $l \to \infty$. This requires for $a = \text{const}$

$$l \sim R^{\alpha}, h \sim R^{\beta} \tag{3.8}$$

with $\alpha, \beta > 0$. Combining this with the previous results we find

$$l = c_1 a^{1-\alpha} R^{\alpha}, \quad h = c_2 a^{1-\beta} R^{\beta} \tag{3.9}$$

with $0 < \alpha, \beta < 1$.

b) This immediately follows from the previous task $l/R = c_1(a/R)^{1-\alpha} \ll 1$.

c) Using the previous tasks we find that the angle θ, see Fig. 3.1, is small: For the angle $\sin(\theta/2) = \frac{l}{2(R+h)}$ with the relation $\theta R + 2\pi a = l$ we find in the limit $R \gg a$ that $\theta \sim l/R$ and, hence, $\theta \to 0$ using the found scaling forms.

We use the following relations. The rope is of length $2\pi(R+a)$ which has to coincide with the length of l and the remaining circle

$$(2\pi - \theta)R + l = 2\pi(R+a) \tag{3.10}$$
$$\Rightarrow l = 2\pi a + \theta R \tag{3.11}$$
$$\Rightarrow \theta = (l - 2\pi a)/R. \tag{3.12}$$

We have shown that θ is small, so that we can write

$$\frac{l}{2R} = \tan(\theta/2) \approx \theta/2 + \frac{1}{3}(\theta/2)^3 . \tag{3.13}$$

This yields

$$\frac{l}{2R} = \frac{l-2\pi a}{2R} + \frac{1}{3}\left(\frac{l-2\pi a}{2R}\right)^3 \tag{3.14}$$
$$\frac{1}{3}\left(\frac{(l-2\pi a}{2R}\right)^3 = \frac{2\pi a}{2R} \tag{3.15}$$
$$\frac{l-2\pi a}{2R} = \left(\frac{3\pi a}{R}\right)^{1/3} \tag{3.16}$$
$$l = 2\pi a + 2R\left(\frac{3\pi a}{R}\right)^{1/3} \to (24\pi)^{1/3} a^{1/3} R^{2/3}. \tag{3.17}$$

Now we obtain h via

$$(R+h)^2 = (l/2)^2 + R^2 \tag{3.18}$$
$$\Rightarrow h = \sqrt{(l/2)^2 + R^2} - R = \frac{l^2}{8R} + \mathcal{O}(R(l/R)^4) \to \frac{(3\pi)^{2/3}}{2} a^{2/3} R^{1/3}. \tag{3.19}$$

d) There are two equations that determine the slope of the rope $y(x)$. The forces acting on a piece of the rope and the reference system are shown in Fig. 3.2. Although solving the equations in the

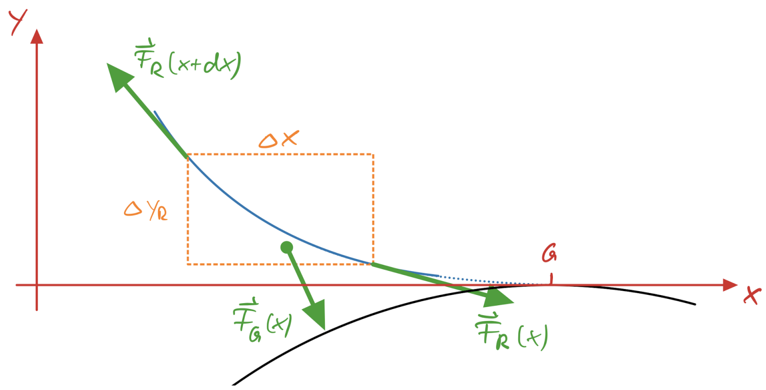


Figure 3.2 Caption

reference system used might be rather complicated, writing down the equations is easier. Note that the gravitational force $g(x, y)$ is position dependent.

One condition stems from the requirement that the rope is in equilibrium and, hence, all forces cancel. We have the tension within the rope $\vec{F}_\mathrm{R}$ and the gravitational force pulling towards the center of the earth

$$\vec{F}_\mathrm{G}(x) = \int_{s(x)}^{s(x+\mathrm{d}x)} \mathrm{d}\tilde{s}\rho\vec{g}(x(\tilde{s}), y(\tilde{s})), \tag{3.20}$$

where we integrated over the length of the rope s from x to $\mathrm{d}x$. We can now substitute x for s by using

$$\frac{\mathrm{d}\,s}{\mathrm{d}x} = \sqrt{1 + \left(\frac{\mathrm{d}\,y}{\mathrm{d}x}\right)^2} \tag{3.21}$$

(heuristically $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$) yielding

$$\vec{F}_\mathrm{G}(x) = \int_x^{x+\mathrm{d}x} \mathrm{d}\tilde{x}\rho\vec{g}(\tilde{x}, y(\tilde{x}))\sqrt{1 + y'(\tilde{x})^2} = \rho\vec{g}(\tilde{x}, y(\tilde{x}))\sqrt{1 + y'(\tilde{x})^2}\mathrm{d}x + \mathcal{O}((\mathrm{d}x)^2). \tag{3.22}$$

Using this we can write down the condition of the cancelling forces

$$\vec{F}_\mathrm{R}(x + \mathrm{d}x) - \vec{F}_\mathrm{R}(x) + \vec{F}_\mathrm{G}(x) = 0 \tag{3.23}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\vec{F}_\mathrm{R}(x) + \vec{F}_\mathrm{G}(x) = 0. \tag{3.24}$$

The second equation needed to determine the slope y is the given by the requirement that in equilibrium, $\vec{F}_\mathrm{R}(x)$ has to be tangential to the rope. This yields

$$\frac{F_{\mathrm{R},y}(x)}{F_{\mathrm{R},x}(x)} = \frac{\mathrm{d}\,y}{\mathrm{d}x}. \tag{3.25}$$

So in total we have three differential equations and three unknown functions $y(x), F_{\mathrm{R},y}(x), F_{\mathrm{R},x}(x)$.

e) As these are first order differential equations, we need three initial conditions to fix the slope. These can be obtained by considering the point G, i.e., where the slope touches the ground. At this point, $F_{\mathrm{R},y}(x = G) = 0$ as the slope is tangential to the earth and $\vec{F}_\mathrm{R}$ is tangential to the rope. This is equivalent to $y'(x = G) = 0$. At $x = G$ we obtain as well $y(x = G) = 0$. The final condition is that the rope is by πa longer than the curvature of the earth $y_E(x)$

$$\int_0^G \mathrm{d}x\,\sqrt{1 + y_E'(x)^2} + \pi a = \int_0^G \mathrm{d}x\,\sqrt{1 + y'(x)^2}. \tag{3.26}$$

So there is only one degree of freedom left, namely G.

f) We simplify the equations by neglecting the curvature of the earth and the spatial dependence of the gravitational force. In that case $h(x) = y(x)$ and $\vec{g} = -g\vec{e}_y$, i.e., the reference system previously chosen is aligned with the earth. The differential equations simplify to

$$\frac{\mathrm{d}}{\mathrm{d}x}F_{\mathrm{R},x}(x) = 0, \tag{3.27}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}F_{\mathrm{R},y}(x) - \rho g\sqrt{1 + y'(x)^2} = 0, \tag{3.28}$$

$$F_{\mathrm{R},y}/F_{\mathrm{R},x} = -y'(x). \tag{3.29}$$

The first yields $F_{\mathrm{R},x}(x) = \mathrm{const} \equiv F_{\mathrm{R},x}$. Using this, the second can be rewritten to

$$F_{\mathrm{R},x} y''(x) - \rho g \sqrt{1 + y'(x)^2} = 0. \tag{3.30}$$

There are various ways to solve this differential equation.

Aside guessing the solution the easiest might be to use the hint with the substitution $x \to y'(x)$ and find

$$\int \mathrm{d}y' \frac{1}{\sqrt{1 + (y')^2}} = \mathrm{asinh}(y') + \mathrm{const} \tag{3.31}$$

$$\Leftrightarrow \mathrm{asinh}(y') = \frac{\rho g}{F_{\mathrm{R},x}} x + C_1. \tag{3.32}$$

Integration of $y'(x)$ yields

$$y(x) = b^{-1} \cosh(bx + C_1) + C_2, \tag{3.33}$$

where we introduced $b = \rho g / F_{\mathrm{R},x}$. We can determine C_1 and C_2 in terms of G via $y(G) = 0, y'(G) = 0$ so that

$$y(x) = b^{-1}(\cosh(b(x - G)) - 1). \tag{3.34}$$

g) We now exchange one constraint of the system in comparison to the previous task: we require $y(0) = \tilde{h} = h/10$, where h is the maximal height given in the first subtask. This fully determines the system.

The aim of this task is to calculate the total tension in the rope at $x = 0$, namely the square root of $F_{\mathrm{R},x}^2 + F_{\mathrm{R},y}^2$.

We obtain the total tension within the rope by finding $F_{\mathrm{R},x}$ by calculating b in dependence of h, a and then $F_{\mathrm{R},y}$ by calculating $y'(0)$ and G in dependence of h, a.

Using the found form, the condition $y(0) = h/10 = \tilde{h}$ translates to

$$(\cosh(bG) - 1)/b = \tilde{h}. \tag{3.35}$$

The condition of a rope of fixed length becomes

$$G + \pi a = \int_0^G \mathrm{d}x \, \sqrt{1 + y'(x)^2} = \int_0^G \mathrm{d}x \, \sqrt{1 + \sinh^2(b(x - G))} = \sinh(bG)/b. \tag{3.36}$$

As given in the task, we use $|y'(0)| \ll 1$. It follows $|bG| \ll 1$. We now have to taylor all expressions up to third order in bG as otherwise $G + \pi a = \sinh(bG)/b$ cannot be fulfilled. Tayloring yields

$$G + \pi a = G + (bG)^3/(6b) \quad \Leftrightarrow \quad \pi a = b^2 G^3/6 \tag{3.37}$$

$$bG^2/2 = \tilde{h}. \tag{3.38}$$

Inserting the two equations into each other allows to deduce

$$\pi a = \left(\frac{2\tilde{h}}{G^2}\right)^2 G^3/6 \tag{3.39}$$

$$\Rightarrow G = \frac{2\tilde{h}^2}{3\pi a} \tag{3.40}$$

$$\Rightarrow b = \frac{2\tilde{h}}{G^2} = \frac{9\pi^2 a^2}{2\tilde{h}^3}. \tag{3.41}$$

Hence, it follows

$$F_{\mathrm{R},x}^2 + F_{\mathrm{R},y}^2 = F_{\mathrm{R},x}^2(1 + [y'(0)]^2) = \left(\frac{4\rho^2 g^2 \tilde{h}^6}{81\pi^4 a^4}\right)\left(1 + \frac{9a^2\pi^2}{\tilde{h}^2}\right). \tag{3.42}$$

We can find the scaling of the solution using the first task so that $a^2/\tilde{h}^2 \ll 1$ and

$$F_{\mathrm{R}}^2 \to \frac{\rho^2 g^2 R^2}{16(10)^6\pi^2}. \tag{3.43}$$

h) The rope breaks.

Problem 4

The Galactic Centre Laboratory

Dr. Odele Straub – *ORIGINS Excellence Cluster and Max Planck Institute for Extraterrestrial Physics*

Background The Milky Way, our home galaxy, is the second largest galaxy (after Andromeda) in the local neighbourhood; it spans about 30 kpc (1 kpc = 1000 pc = 3.1 $\times 10^{19}$ m) in diameter. Deep in its centre resides the supermassive, compact radio source, Sagittarius A* (Sgr A*). While most bright and young stars in the Milky Way are located in the gas rich spiral arms of the galactic disc, the overall star count increases towards the centre. In particular in the innermost 0.04 pc there is a dense cluster of about 100 young and fast moving stars called the S-stars. Their orbits around the central gravitating mass have random orientations. This central region cannot be observed at optical wavelengths due to the thick molecular clouds in our line of sight. However, in the infrared it is possible to pierce through the dust. GRAVITY is a beam-combiner instrument that links the four 8-meter infrared telescopes of the Very Large Telescope (VLT) in Chile into one giant telescope with a diameter of $D = 130$ m. With its high angular resolution, GRAVITY can monitor extremely faint and distant objects, like the stars in the immediate vicinity of Sgr A* , with a precision that allows to record daily changes in their motion. Consequently, the Galactic centre region has now become a new "laboratory" to probe and test general relativity.

Useful constants: the gravitational constant, G = 6.7 $\times 10^{-11} \mathrm{m}^3/\mathrm{kg\ s}^2$, the speed of light in vacuum, c = 3 $\times 10^8$ m/s.

a) [*2 points*] **Mass of the Milky Way:** The Sun with its mass of $M_\odot = 1.99 \times 10^{30}$ kg is an *average star*. It sits at the edge of a spiral arm of the Milky Way and orbits Sgr A* at a distance of $r_0 = 8.28$ kpc with a circular velocity of $v_\odot = 251.05$ km/s.

(i) Estimate the mass of the Milky Way in units of solar masses.

(ii) Explain why this is only a lower limit.

The actual stellar content can be derived from luminosity measurements and amounts to roughly 15% of the total mass of the Milky Way.

(iii) Where and/or what is the rest?

b) [*2 points*] **Mass of Sgr A* :** Astronomers deduce the mass of Sgr A* from the motion of the S-stars. The star S2 (see Fig. 4.1) is particularly well suited due to its short orbital period of $P = 16.05$ years, small semi-major axis of $a = 0.125$" (arcseconds) and high eccentricity $e = 0.88$. Its closest approach to the central gravitating mass, i.e. its pericentre is $r_{\mathrm{peri}} = 14$ mas (milli-arcseconds). First convert the angular size of the orbit from arcseconds to SI units with the help of the Sun's distance from Sgr A* . Then calculate the mass of Sgr A* in units of solar masses.

c) [*2 points*] **Size of Sgr A* :** A black hole (BH) is a mathematical object native to a theory of gravity. It is defined by having a horizon instead of a surface, i.e. a critical radius from where not even light can *escape*.

(i) Derive the critical radius from Newtonian principles. The resulting formula is also found in the theory of General Relativity (GR) and called Schwarzschild radius, r_S.

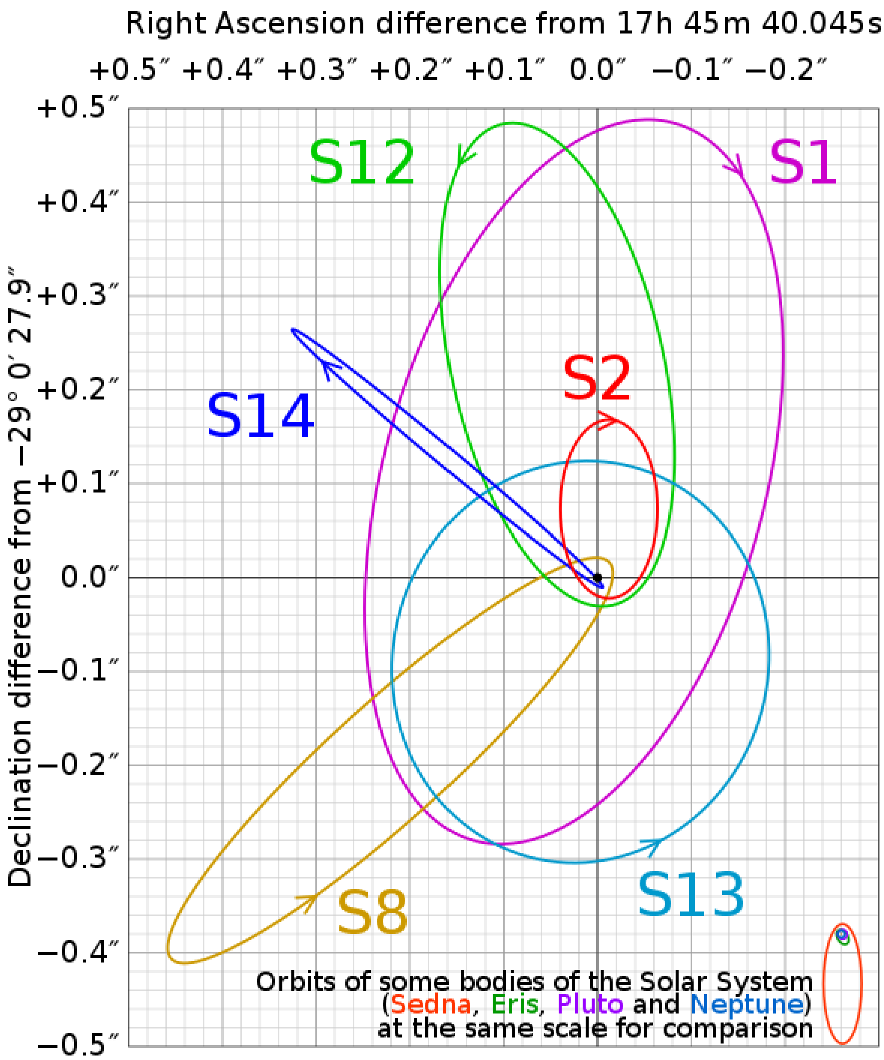


Figure 4.1 Orbits of some of the inner S-stars around Sgr A* , the supermassive compact object in the centre of the Milky Way. The star S2 (red) is due to its short and highly eccentric orbit the most interesting probe of the gravitational field of Sgr A* . At the bottom right, for comparison, some orbits of Solar System bodies. Figure credit: Eisenhauer et al. (2005)

(ii) Calculate r_S of Sgr A* for the mass you determined above and compare it to Neptune's mean distance from the Sun, $r_{\text{Neptune}} = 4.5$ billion km.

The GRAVITY/VLT instrument not only sees the stars near Sgr A* but also detects flickering light from a location even closer. This light originates from occasional hot plasma flares. They loop around Sgr A* with an average radius of 60 μas (micro-arcseconds).

(iii) Argue, considering theory and observations, why Sgr A* must be a compact object and is likely to be a BH.

d) [*1 point*] **GRAVITY Instrument:** The perhaps most important equation in (observational) astronomy gives the angular resolution R of any telescope in units of radians for any given wavelength

λ and telescope diameter D

$$R = 1.22\frac{\lambda}{D}, \tag{4.1}$$

GRAVITY/VLT observes at infrared wavelengths of 2.2 μm. What is its resolution? Estimate the diameter of an infrared telescope needed to resolve the black hole horizon. Where would you build it?

e) [*1 point*] **Gravitational Redshift:** There are three classical tests of GR proposed by Albert Einstein: the perihelion precession of planet Mercury, the deflection of light by the Sun, and the gravitational redshift of light. To detect the effect of gravitational redshift in the galactic centre, one follows the star S2 on its orbit. Astronomers not only track its positions (with GRAVITY) but also record the radial velocities using a spectrometer (e.g. SINFONI, or ERIS at the VLT). The stellar spectrum shows a prominent absorption line at a wavelength λ'. Explain what gravitational redshift is and how it is related to the star's velocity. Where do you expect the strongest effect during the orbit of S2?

f) [*2 points*] **Precession of the Pericentre:** Recently, astronomers showed that the star S2 precesses around Sgr A* . That is to say, S2 is not moving on a closed ellipse, but on an open rosette-like trajectory. With this finding GR passed another test in the Galactic centre. S2 currently approaches the apocentre of its orbit, i.e. the farthest point from Sgr A* . Observations indicate that the pericentre of S2 advances each orbit by a small angle, $\delta\varphi = 12'$/orbit. Calculate by how much the position of the apocentre changes (in mas). Can GRAVITY resolve this? Note: with the help of adaptive optics the exposure time on a target star can be prolonged so that the actual resolution is about ten times better than the nominal R-value you calculated above.

Solution

a) The solar velocity is given by the local standard of rest in Reid et al (2020), $\Theta = 30.32$ km/s/kpc. With the GRAVITY Col. (2021) measurement of $r_0 = 8.28$ kpc this gives an angular speed of $v_\odot = 251.05$ km/s. For any body on a stable circular orbit, the centripetal and the gravitational force are in equilibrium. One can write

$$\frac{m_{\text{S2}} v_\odot^2}{r_0} = \frac{G m_{\text{S2}} M_{\text{enclosed}}}{r_0^2} \tag{4.2}$$

$$M_{\text{enclosed}} = \frac{v_\odot^2 r_0}{G} \tag{4.3}$$

$$= 2.41 \times 10^{41}\,\text{kg}, \tag{4.4}$$

where $r_0 = 8.28\,\text{kpc} = 2.55 \times 10^{20}$ m. Alternatively, one can calculate the orbital period of the Sun from the distance to Sgr A* and the solar circular velocity

$$P_\odot = 2\,\pi r_0 / V_\odot = 6.415 \times 10^{15} s. \tag{4.5}$$

It takes the Sun a bit more than 200 million years to go once around the Galactic centre. Then, relating the orbital velocity to the period via

$$v = \frac{2\pi r}{P} \tag{4.6}$$

one finds Kepler's 3rd Law

$$M_{\text{enclosed}} = \left(\frac{2\pi}{P_\odot}\right)^2 \frac{r_0^3}{G} = 2.41 \times 10^{41}\,\text{kg} \tag{4.7}$$

The enclosed mass given in solar masses is then $\underline{\underline{M_{\text{enclosed}} = 1.21 \times 10^{11}\ M_\odot}}$.

This is a lower mass estimate on the mass of the Milky Way. The Sun at r_0 from the centre is located somewhere near the middle of the Milky Way disc. The calculation does not account for the stars outside the Sun's orbit. Recent calculations (e.g. Watkins et al. 2019, using Gaia DR2 and HST measurements of the proper motion of globular clusters) find a mass of about $4 \times 10^{11}\ M_\odot$. This value is associated to the stellar disc with an outer radius of 39.5 kpc. The calculated mass is not the pure stellar mass content but the total mass, i.e. the stellar/baryonic matter as well as the dark matter content. The stellar mass content is only about 15% of the total mass of the Milky Way. The dark matter envelopes the entire Milky Way in a halo. It has a density distribution that is mostly flat but increases towards the Galactic centre.

b) First convert the semi-major axis of S2 from arc-seconds to metres using r_0

$$arcsec = \frac{1}{3600}\frac{\pi}{180} \cdot r_0 = 1.247 \times 10^{15}\,\text{m}. \tag{4.8}$$

The semi-major axis of S2 is $a = 1.560 \times 10^{14}$ m. Then use Kepler's third law Eq.4.7 to calculate the mass of Sgr A* in SI units

$$M_{SgrA^*} = \left(\frac{2\pi}{P_{\text{S2}}}\right)^2 \frac{a^3}{G} = 8.762 \times 10^{36}\,\text{kg} \tag{4.9}$$

which gives, divided by the mass of the Sun, a black hole mass of

$$\underline{\underline{M_{SgrA^*} = 4.381 \times 10^6\, M_\odot}}. \tag{4.10}$$

c) The Newtonian escape velocity of any gravitating body is derived by equating the kinetic and gravitational energies

$$E_{\text{kin}} = E_{\text{pot}} \tag{4.11}$$
$$\frac{1}{2}mv^2 = \frac{GMm}{r} \tag{4.12}$$
$$\sqrt{\frac{2GM}{r}} = v_{\text{escape}} \tag{4.13}$$

Light (i.e. c, the speed of light) does not escape black holes. The critical radius is then given by setting $v_{\text{escape}} = c$ so that

$$r_{\text{crit}} = \frac{2GM}{c^2} = r_S \tag{4.14}$$

The critical, or Schwarzschild radius of Sgr A* one finds by setting $M = M_{SgrA^*}$ which gives

$$\underline{\underline{r_S = 1.26 \times 10^{10} m}}. \tag{4.15}$$

The observation of flares indicate that an invisible object resides inside a radius of 60 μas. With the conversion

$$1" = \frac{1}{3600}\frac{\pi}{180} r_0 = 1.24 \times 10^{15}\, m \tag{4.16}$$

the flare radius is

$$r_{\text{flare}} = 60\, \mu as = 7.42 \times 10^{10}\, m \tag{4.17}$$

This implies that the flares orbit around Sgr A* at about 6 r_S. The entire mass of Sgr A* must be inside this radius.

A star of the same mass (assumeing an average density of about $\propto 10^3 kg/m^3$) would not fit in. A more compact object like a neutron star (average density of about $\propto 10^{17} kg/m^3$) might. However, in neutrons stars, the Fermi pressure that arises due to degenerated atomic nuclei and this can support masses up to about 2.5 - 3 $M_\odot$ (the precise maximal value depends on the equation of state of the neutron star interior and is to date unknown). There is no known mechanism that supports masses greater than 3 $M_\odot$. For all intents and purposes Sgr A* is considered to be a massive black hole.

d) Using $D = 130$ m from the introduction and $\lambda = 2.2\, \mu$m one obtains

$$R = 1.22\frac{\lambda}{D} = 3.69\,\text{mas}. \tag{4.18}$$

In order to resolve the radius of Sgr A* calculated above, now given in μas

$$r_S = 10.21\, \mu as, \tag{4.19}$$

one has to convert it to radians (divide the above r_S by (180/Pi · 3600)) to find the required telescope diameter of

$$D = 1.22\frac{\lambda}{r_S(180/Pi * 3600)} = \underline{\underline{54.22\, km}} \tag{4.20}$$

e) Gravitational redshift is the phenomenon that photons travelling out of a gravitational well lose energy. Given that the speed of light is constant and the photon cannot slow down, this loss of energy corresponds to an increase in wavelength (or decrease in frequency) towards the redder part of the spectrum. The radial velocity is derived from the Doppler shift of the wavelength λ' of the Br_γ absorption line seen in the spectrum of the moving S2 and compared to the value at rest (λ)

$$\frac{\lambda' - \lambda}{\lambda} = \frac{v_\perp}{c}. \tag{4.21}$$

The total gravitational redshift is composed of a special relativistic component, the transverse Doppler shift, and a general relativistic component, the actual gravitational redshift. The effect is strongest when S2 is closest to Sgr A* , at the pericentre.

f) We know that the pericentre advances after one orbit (after 16 years) by

$$\delta\varphi_{peri} = 6\pi \frac{GM}{c^2 a(1-e^2)} = 12.1'/\text{orbit} \tag{4.22}$$

in one direction (to the east for S2). This implies the apocentre advances about the focal point (Sgr A*) by the same angle to the other direction. The problem can be solved by trigonometry. First, calculate the distance from Sgr A* to the apocentre

$$r_{\text{apo}} = 2 \cdot a - r_{\text{peri}} = 236\,\text{mas}. \tag{4.23}$$

Then, consider the isosceles triangle $\{r_{\text{apo}\,1} \rightarrow Sgr\ A^* \rightarrow r_{\text{apo}\,2}\}$ where $\delta\varphi_{peri}$ is the angle between the two apo-legs. Split the angle in half and calculate the half-separation between the apocentres, then multiply by 2 to the total separation

$$x = 2 \cdot r_{\text{apo}} \cdot \sin\left(\frac{\delta\varphi}{2}\frac{\pi}{180}\right) = \underline{\underline{0.83\,\text{mas}}}. \tag{4.24}$$

Problem 5

Conditions for a Self-Heated Fusion Plasma

Prof. Dr. Sibylle Günter – *Max Planck Institute for Plasma Physics, Garching bei München, Germany*

Background In a fusion power plant, energy shall be released by the fusion reaction of a deuterium and a tritium nuclei to an α particle and a neutron:

$$\mathrm{D} + \mathrm{T} \rightarrow {}^4\mathrm{He} + n + 17.5\mathrm{MeV}. \tag{5.1}$$

To overcome the Coulomb barrier, the reactants need a sufficient kinetic energy. The number of Coulomb collisions is however always larger than the number of fusion reactions. Therefore, a positive energy balance can only be achieved in a thermal plasma. The D-T reaction has the highest fusion rate compared to any other reaction, at lowest plasma temperature. Nevertheless, a plasma temperature of about 10 keV is required. In a future reactor, the plasma should be heated by the energy released in the fusion reactions. The neutrons leave the plasma nearly without interactions, and thus only the α particles contribute to plasma heating. The heat (=energy) transport of a magnetically confined fusion plasma is determined by several effects, of particular importance are radiation (mostly bremsstrahlung) and turbulent transport. To characterize the energy losses, often the so-called energy confinement time τ_E is used, a measure that corresponds to the time after which the plasma is significantly cooled down (after heating is switched off).

a) [*2 points*] Calculate how the total energy of 17.5 MeV is divided between ^{4}He and the neutron in the centre of mass frame of the fusing particles.
Hint: *Use momentum and energy balance.*

b) [*2 points*] Calculate the heating power P_{heat} due to the α particles for a fusion power plant (volume: $V = 1000\,\mathrm{m}^3$) with a constant electron density of $n_e = 10^{20}\,\mathrm{m}^{-3}$. Assume that the plasma consists of 50% D and 50% T ions and the electrons. Keep in mind that the plasma is always (quasi-)neutral, i.e. the number of electrons balances the number of ions. For a thermal plasma, the reactivity (number of fusion reactions per volume per time) is approximately given by $\langle\sigma v\rangle \approx 10^{-22}\,\mathrm{m}^3\mathrm{s}^{-1}$ for $T = 10\,\mathrm{keV}$.

c) [*2 points*] Provide an expression for the loss power. The loss power P_{loss} is defined by the thermal energy of the plasma devided by the energy confinement time τ_E. Assume that the plasma behaves like an ideal gas.

d) [*2 points*] Balance plasma heating and loss power to derive the so-called Lawson criterion, a criterion in terms of plasma density $n = n_i$, i.e. equal to the ion density, temperature T and energy confinement time τ_E. Please note: In the temperature range considered (≈10 keV) the fusion reactivity increases with temperature approximately proportional to T^2.

e) [*2 points*] As the fusion reactions produce ^{4}He, it is not consistent to assume that the plasma consists of D and T only. By how much would the fusion power be reduced as compared to a pure D-T plasma if 10% of the plasma ions would be ^{4}He?

Solution

a) As momentum and energy of the fusion products are much higher than that of D-T, the momentum balance can be written as:

$$m_\alpha v_\alpha \approx -m_n v_n. \tag{5.2}$$

Energy conservation is given by

$$\frac{m_\alpha}{2} v_\alpha^2 + \frac{m_n}{2} v_n^2 \approx Q_{DT} \approx 17.5\,\text{MeV}. \tag{5.3}$$

From these equations one finds

$$\frac{E_\alpha}{E_n} = \frac{m_n}{m_\alpha} = \frac{1}{4} \tag{5.4}$$

and thus $E_n = 14.1$ MeV and $E_\alpha = 3.5$ MeV.
0.5 points for correct momentum balance
0.5 points for correct energy balance
0.5 points for energy ratio is mass ratio
0.5 points for correct results

b) The fusion power available for plasma heating is given by

$$P_{heat} = n_D n_T \langle \sigma v \rangle E_\alpha V. \tag{5.5}$$

As the plasma consists of 50% D and 50% T, this results in

$$P_{heat} = \frac{n^2}{4} \langle \sigma v \rangle E_\alpha V, \tag{5.6}$$

where $n = n_e = n_i$ is the plasma density equal to electron and ion density in a (quasi-)neutral plasma and we used that $n_D = n_T = n/2$. With $E_\alpha = 3.5\,\text{MeV}$ and $V = 1000\,\text{m}^3$ one finds

$$P_{heat} \approx 140\text{MW}. \tag{5.7}$$

1.5 points for correct deviation
0.5 points for correct numerical result

c) Considering the plasma as an ideal gas, the plasma energy is given by

$$E_{plasma} = \frac{3}{2}(n_e + n_i) k_B T V, \tag{5.8}$$

where n_e is the electron density and n_i is the ion density. Assuming quasi neutrality ($n_e = n_i = n$) the loss power is given by

$$P_{loss} = \frac{E_{plasma}}{\tau_E} = \frac{3 n k_B T}{\tau_E} V. \tag{5.9}$$

1.5 points for correct equation for energy
0.5 points for correct equation of P_{loss}

d) Assuming $\langle \sigma v \rangle = c_r T^2$ and balancing the heat power

$$P_{heat} = \frac{n_D n_T}{2} \langle \sigma v \rangle E_\alpha V \approx \frac{n^2}{4} c_r T^2 E_\alpha V \tag{5.10}$$

with the loss power

$$P_{loss} = \frac{3nk_BT}{\tau_E}V \tag{5.11}$$

gives the Lawson criterion:

$$nT\tau_E > \frac{12k_B}{c_r V E_\alpha} =: c. \tag{5.12}$$

(Given the numbers above $c \approx 3 \cdot 10^{21}$keV s/m^3).
0.5 points for correct $\langle \sigma v \rangle$ dependence
0.5 points for correct balancing ansatz
1 point for correct Lawson-criterion

e) 10% He provides 20% of the plasma electrons (as Z=2) and thus reduces the ion density $n_i = n_D + n_T$ to 80%. As the fusion power is proportional to $n_i^2 = (n_D + n_T)^2$, the fusion power would be reduced by 36%.
1 point for correct estimation of ion density
1 points for correct result
Alternative argumentations are fine if correct

Problem 6

Boltzmann-Factors from Information Entropy

Prof. Dr. Björn Malte Schäfer – *Fakultät für Physik und Astronomie, Heidelberg University, Germany*

Background Statistical mechanics operates under the assumption (called the fundamental postulate) that in thermal equilibrium all states at a given energy are equally likely (defining the microcanonical ensemble) and if energy can fluctuate, states with energy difference $\Delta\epsilon$ are populated according to the Boltzmann factor,

$$p(\Delta\epsilon) = \exp\left(-\frac{\Delta\epsilon}{k_B T}\right), \tag{6.1}$$

T being the temperature and k_B the Boltzmann-constant. There is a more fundamental idea though: Claude Shannon has shown that the information entropy S

$$S = -\int \mathrm{d}x\, p(x) \ln p(x) = -\langle \ln p \rangle \tag{6.2}$$

is a positive measure of randomness of a distribution $p(x)$ and is additive for independent random processes. There are also alternatives, for instance the entropy measure,

$$S_\alpha = \frac{1}{1-\alpha} \ln \int \mathrm{d}x\, p^\alpha(x) = -\frac{1}{\alpha - 1} \ln\langle p^{\alpha-1} \rangle \tag{6.3}$$

with a free positive parameter $\alpha \neq 1$ proposed by Alfred Rényi.
One could now try to reason like this: (i) thermal equilibrium should correspond to the state of highest randomness, and as (ii) information entropy is such a measure of randomness, the distribution of systems of an ensemble in thermal equilibrium should realize the highest information entropy.

Information entropies S and S_α are functionals for the actual distribution $p(x)\mathrm{d}x$, x being the collection of phase space coordinates, and therefore one can carry out functional variations and find the corresponding distributions: We shall try this for Shannon's entropy as well as for Rényi's entropy!

a) [*2 points*] Let's get ready with entropies.
Please compute the entropy S for a Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{6.4}$$

as a function of σ^2. What is the Shannon-entropy S for a step distribution

$$p(x) = \begin{cases} \frac{1}{b-c} & \text{if } c < x < b \\ 0 & \text{else} \end{cases}, \tag{6.5}$$

does the entropy increase if the interval or the variance become larger? What is the physical interpretation of this? Now, try out the Rényi-entropy for both distributions: Do they scale in a similar way with σ^2 or $b - a$?

<u>Hint</u>: *The relation* $\int \mathrm{d}x\, e^{-x^2} = \sqrt{\pi}$ *might be useful.*

b) [*2 points*] Micro- and canonical ensembles from Shannon's entropy
The constant distribution for the microcanonical ensemble is derived like this: One determines the distribution that maximises S under the boundary condition $\int \mathrm{d}x\, p(x) = 1$. This leads to the following variation

$$\delta\left[S(p) + \lambda\left(\int \mathrm{d}x\, p(x) - 1\right)\right] = -\delta \int \mathrm{d}x\, p(x) \ln p(x) + \lambda\delta\left(\int \mathrm{d}x\, p(x) - 1\right) = 0 \tag{6.6}$$

with a Lagrange-multiplier λ. The solution is $p(x) = e^{(\lambda-1)} = \mathrm{const}$, where λ can in principle be determined from $\int \mathrm{d}x\, p(x) = 1$. A constant distribution maximises therefore S, and the constant distribution of systems across phase space x would already be the microcanonical ensemble! Please solve equation (6.6) and confirm that $p(x) = e^{(\lambda-1)} = \mathrm{const}$ is a solution.

Let's try out to get the Boltzmann-factor: Please derive the distribution $p(x)$ that maximises S with the additional boundary condition $\int \mathrm{d}x\, p(x)\epsilon(x) = \epsilon$ (imposed with a second Lagrange-multiplier μ), here $\epsilon(x)$ is a function which returns the energy for given system-coordinates x, the exact form of this function is irrelevant for the exercise. Furthermore, show that

$$\frac{p(\epsilon(x_2))}{p(\epsilon(x_1))} = \exp\left(\mu \cdot (\epsilon(x_2) - \epsilon(x_1))\right) \tag{6.7}$$

at fixed x, which has already the shape of a Boltzmann-factor. Please show by using the definition of temperature T in the microcanonical ensemble,

$$\frac{\partial S}{\partial \epsilon} = \frac{1}{k_B T} \tag{6.8}$$

that the Lagrange-multiplier needs to be $\mu = -1/(k_B T)$.

Hint: *Here is a short recap on variational calculus: You may view the entropy S as a functional, mapping a function $p : \mathbb{R} \to \mathbb{R}$ to a real value $S(p) = -\int \mathrm{d}x\, p(x) \ln p(x)$. For such a functional, the first variation is given by $\delta S(p) = \frac{\partial}{\partial \epsilon} S(p + \epsilon h)|_{\epsilon=0}$, where h is another function from the same vector space as p. If the first variation of S around p equals 0, you know that p maximizes or minimizes S. The same method can also be used with boundary conditions by adding them via constant Lagrange-multipliers, as shown in equation (6.6).*
For this task, you don't have to determine the exact value of the Lagrange-multiplier λ.

c) [*2 points*] Micro- and canonical ensembles from Rényi's entropy
Show that the constant distribution maximises the Rényi entropy S_α, which would correspond to the microcanonical ensemble. Generalising this result with a constraint on energy, can you derive the ratio $\frac{p(\epsilon(x_2))}{p(\epsilon(x_1))}$? What needs to hold for $\epsilon(x_2) - \epsilon(x_1)$, such that you can approximate the result to a similar form as the result for the Shannon entropy. At last derive the Boltzmann-factor μ from the maximised Rényi entropy and $\frac{\partial S}{\partial \epsilon} = \frac{1}{k_B T}$ (without any approximations).

Hint: *Again, the actual value of the Lagrange-multiplier λ does not matter.*

d) [*2 points*] Equivalence of Shannon and Rényi-entropies
Please show that in the limit $\alpha \to 1$ one recovers Shannon's entropy measure from the Rényi-entropy,

$$\lim_{\alpha \to 1} S_\alpha = S \tag{6.9}$$

Hint: *Use de l'Hôpital's rule.*

e) [*2 points*] Choice of Shannon's entropy
The familiar Boltzmann-factor comes out if one chooses Shannon's entropy as a measure of randomness, and Rényi's entropy would not reproduce it: What is it in the mathematical formulation about the Rényi-entropy that makes it contradictory to physical observations?

Hint: *That's a tough question and there are different strategies to answer it. Maybe think about the case of multidimensional system and conditional probabilities. Alternatively it might help to argue with the scaling of a physical law in mind.*

Solution

a) Let's calculate the entropy for a Gaussian distribution

$$
\begin{aligned}
S =& -\int \mathrm{d}x\, p(x) \ln\left[p(x)\right] \\
S =& -\int \mathrm{d}x\, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \left(-\frac{1}{2}\ln\left[2\pi\sigma^2\right]\right) \\
& -\int \mathrm{d}x\, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \left(\frac{-x^2}{2\sigma^2}\right) \\
=& \frac{1}{2}\ln\left[2\pi\sigma^2\right] + \frac{1}{2}\ln\left[e\right] \\
=& \frac{1}{2}\ln\left[2\pi\sigma^2 e\right]. \quad [0.5 \text{ Points}]
\end{aligned}
$$

Here, we used $\int \mathrm{d}x\, e^{-x^2} = \sqrt{\pi}$.

Now the calculation for the constant function

$$
\begin{aligned}
S =& -\int \mathrm{d}x\, p(x) \ln\left[p(x)\right] \\
=& -\int_c^b \mathrm{d}x\, \frac{1}{b-c} \ln\left[\frac{1}{b-c}\right] \\
=& \ln\left[b-c\right]. \quad [0.5 \text{ Points}]
\end{aligned}
$$

The Shannon-entropies of both distributions increase logarithmically with σ^2 and $(b-c)$, respectively. If σ^2 and $(b-c)$ are small, the state of the system is fairly localized, leading to a small entropy, since small entropy goes hand in hand with a small uncertainty in the system.
Now let's have a look at the Rényi-entropies:

$$
\begin{aligned}
S_\alpha =& -\frac{1}{\alpha-1} \ln \int \mathrm{d}x\, p^\alpha(x) \\
=& -\frac{1}{\alpha-1} \ln\left[\int \mathrm{d}x \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^\alpha e^{-\frac{\alpha x^2}{2\sigma^2}}\right] \qquad y = \sqrt{\alpha}x; \quad \mathrm{d}y = \sqrt{\alpha}\mathrm{d}x \\
=& -\frac{1}{\alpha-1} \ln\left[\int \mathrm{d}y \frac{1}{\sqrt{\alpha}} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{(\alpha-1)} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{-\frac{y^2}{2\sigma^2}}\right] \\
=& \frac{1}{\alpha-1} \ln\left[\sqrt{\alpha\left(2\pi\sigma^2\right)^{(\alpha-1)}}\right]. \quad [0.5 \text{ Points}]
\end{aligned}
$$

Now for the linear distribution:

$$
\begin{aligned}
S_\alpha =& -\frac{1}{\alpha-1} \ln \int \mathrm{d}x\, p^\alpha(x) \\
=& -\frac{1}{\alpha-1} \ln\left[\int_c^b \mathrm{d}x \left(\frac{1}{b-c}\right)^\alpha\right] \\
=& \frac{1}{1-\alpha} \ln\left[(b-c)^{(1-\alpha)}\right] \\
=& \ln\left[b-c\right]. \quad [0.5 \text{ Points}]
\end{aligned}
$$

The Rényi-entropies for the two distributions have the same scaling as the Shannon entropies! [If all integrals are right, but the scaling was not compared substract 0.5points]

b) Prove that a constant function maximizes the entropy:

$$\delta S = -\delta \int \mathrm{d}x\, p(x) \ln p(x) + \lambda \delta \left(\int \mathrm{d}x\, p(x) - 1 \right) = 0$$

$$\delta S = \int \mathrm{d}x\, \left[- \ln p(x) - 1 + \lambda \right] \delta p = 0$$

$$\rightarrow \ln p(x) = \lambda - 1$$

$$\rightarrow p(x) = e^{\lambda - 1} \quad [0.5 \text{ Points}]$$

Same strategy for the Boltzmann-factor:

$$\delta S = -\delta \int \mathrm{d}x\, p(x) \ln p(x) + \lambda \delta \left(\int \mathrm{d}x\, p(x) - 1 \right) + \mu \delta \left(\int \mathrm{d}x\, \epsilon(x) p(x) - \epsilon \right) = 0$$

$$\rightarrow \ln p(x) = \lambda + \mu\epsilon - 1$$

$$\rightarrow p(x) = e^{\lambda + \mu\epsilon - 1}. \quad [0.5 \text{ Points}]$$

Therefore the ratio at two different energies $\epsilon_1 = \epsilon(x_1)$ and $\epsilon_2 = \epsilon(x_2)$ is

$$\frac{p_1}{p_2} = e^{(\mu(\epsilon_1 - \epsilon_2))} \quad [0.5 \text{ Points}]$$

Now we have to determine the prefactor μ via $\frac{\partial S}{\partial \epsilon} = \frac{1}{k_B T}$:

$$\begin{aligned} \frac{\partial S}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} \left[- \int \mathrm{d}x\, p(x) \ln \left(e^{\lambda + \mu\epsilon(x) - 1} \right) \right] \\ &= \frac{\partial}{\partial \epsilon} \left[- \int \mathrm{d}x\, p(x) \left(\lambda + \mu\epsilon(x) - 1 \right) \right] \\ &= -\frac{\partial}{\partial \epsilon} \left[\lambda + \mu\epsilon - 1 \right] = -\mu \\ \rightarrow \mu &= -\frac{1}{k_B T}. \quad [0.5 \text{ Points}] \end{aligned}$$

c) We are repeating the calculation for the Rényi entropy. First the constant distribution:

$$\begin{aligned} \delta S &= \frac{-1}{\alpha - 1} \delta \ln \left[\int \mathrm{d}x\, p(x)^\alpha \right] + \lambda \delta \left(\int \mathrm{d}x\, p(x) - 1 \right) = 0 \\ &= \int \mathrm{d}x\, \left[\frac{-\alpha}{\alpha - 1} \frac{1}{\int \mathrm{d}x\, p^\alpha} p^{\alpha - 1} + \lambda \right] \delta p \\ \rightarrow p(x) &= \left(\lambda \frac{\alpha - 1}{\alpha} C \right)^{\frac{1}{\alpha - 1}} \quad \text{with} \quad C = \int \mathrm{d}\, x p(x)^\alpha. \quad [0.5 \text{ Points}] \end{aligned}$$

Now let's calculate the Boltzmann factor:

$$\begin{aligned} \delta S &= \frac{-1}{\alpha - 1} \ln \left[\delta \int \mathrm{d}x\, p(x)^\alpha \right] + \lambda \delta \left(\int \mathrm{d}x\, p(x) - 1 \right) + \mu \delta \left(\int \mathrm{d}x\, \epsilon(x) p(x) - \epsilon \right) = 0 \\ &= \int \mathrm{d}x\, \left[\frac{-\alpha}{\alpha - 1} \frac{1}{\int \mathrm{d}x\, p^\alpha} p^{\alpha - 1} + \lambda + \mu\epsilon(x) \right] \delta p \\ \rightarrow p(x) &= \left((\lambda + \mu\epsilon(x)) \frac{\alpha - 1}{\alpha} C \right)^{\frac{1}{\alpha - 1}} \quad \text{with} \quad C = \int \mathrm{d}\, x p(x)^\alpha. \quad [0.5 \text{ Points}] \end{aligned}$$

We obtain for the ratio at two different energies:

$$\begin{aligned}\frac{p(\epsilon_2)}{p(\epsilon_1)} &= \left(\frac{\lambda+\epsilon_2\mu}{\lambda+\epsilon_1\mu}\right)^{\frac{1}{\alpha-1}}\\ &= e^{\left[\frac{1}{\alpha-1}\ln\left(\frac{\lambda+\epsilon_2\mu}{\lambda+\epsilon_1\mu}\right)\right]}\\ &= e^{\left[\frac{1}{\alpha-1}\ln\left(\frac{1+\epsilon_2\frac{\mu}{\lambda}}{1+\epsilon_1\frac{\mu}{\lambda}}\right)\right]}\\ &\approx e^{\left[\frac{1}{\alpha-1}\frac{\mu}{\lambda+\mu\epsilon_1}(\epsilon_2-\epsilon_1)\right]} \quad \text{for small energy difference.} \quad [0.5 \text{ Points}]\end{aligned}$$

The Rényi entropy gives only in approximation a normal exponential scaling! Now we calculate μ.

$$\begin{aligned}\frac{\partial S}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon}\frac{-1}{\alpha-1}\ln\left[\int \mathrm{d}x p(x)^{\alpha-1}p(x)\right]\\ &= \frac{\partial}{\partial \epsilon}\frac{-1}{\alpha-1}\ln\left[\int \mathrm{d}x\left((\lambda+\mu\epsilon(x))\frac{\alpha-1}{\alpha}C\right)p(x)\right]\\ &= \frac{\partial}{\partial \epsilon}\frac{-1}{\alpha-1}\ln\left[\left((\lambda+\mu\epsilon)\frac{\alpha-1}{\alpha}C\right)\right]\\ \frac{1}{k_BT} &= \frac{-1}{\alpha-1}\frac{1}{(\lambda+\mu\epsilon)}\mu\\ \frac{\lambda+\mu\epsilon}{k_BT} &= \frac{-1}{(\alpha-1)}\mu\\ \mu &= \frac{-\lambda}{k_BT}\frac{1}{\frac{1}{(\alpha-1)}+\frac{\epsilon}{k_BT}} \quad [0.5 \text{ Points}]\end{aligned}$$

d) We want to prove:

$$\lim_{\alpha\to1} S_\alpha = S$$

We use the de l'Hôpital's rule $\lim\limits_{x\to1}\frac{f(x)}{g(x)} = \lim\limits_{x\to1}\frac{f'(x)}{g'(x)}$ with $f(x) = -\ln\left[\int \mathrm{d}x\, p(x)^\alpha\right]$ and $g(x) = \alpha - 1$. We obtain:

$$\begin{aligned}\lim_{\alpha\to1} S_\alpha &= \lim_{\alpha\to1}\frac{\frac{\partial}{\partial\alpha}\int \mathrm{d}x\, p(x)^\alpha}{\int \mathrm{d}x\, p(x)^\alpha} \quad [0.5 \text{ Points}]\\ &= -\lim_{\alpha\to1}\frac{\int \mathrm{d}x\, \frac{\partial}{\partial\alpha}e^{\ln(p(x)^\alpha)}}{\int \mathrm{d}x\, p(x)^\alpha}\\ &= -\lim_{\alpha\to1}\frac{\int \mathrm{d}x\, \frac{\partial}{\partial\alpha}e^{\alpha\ln(p(x))}}{\int \mathrm{d}x\, p(x)^\alpha} \quad [0.5 \text{ Points}]\\ &= -\lim_{\alpha\to1}\frac{\int \mathrm{d}x\, \ln(p(x))\, e^{\alpha\ln(p(x))}}{\int \mathrm{d}x\, p(x)^\alpha} \quad [0.5 \text{ Points}]\\ &= -\lim_{\alpha\to1}\frac{\int \mathrm{d}x\, \ln(p(x))\, p(x)^\alpha}{\int \mathrm{d}x\, p(x)^\alpha}\\ &= \int \mathrm{d}x\, \ln(p(x))\, p(x) \quad [0.5 \text{ Points}]\end{aligned}$$

e) Really difficult to answer. All creative and correct thoughts will get 2 points.
One idea: Thermodynamics is a theory of information about statistical systems, and Shannon's entropy is the only one to fulfil the law of conditional probability $S(y|x) + S(x) = S(x, y)$ and is therefore compatible with Bayes' law $p(y|x) \cdot p(x) = p(x, y)$ with $S = \langle \ln p \rangle$. Another idea would be to argue, that thermodynamic laws of nature show an exponential scaling and not a power law scaling.

Problem 7

Active Brownian Particle

Prof. Dr. Michael Schmiedeberg – *Institut für Theoretische Physik 1, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany*

Background Self-propelled particles, also known as active particles, are popular as a model system for an intrinsically non-equilibrium system in statistical physics. Furthermore, they are used to mimic biological or socio-economic systems like colonies of swimming bacteria, fish schools, animal flocks, swarms of birds or insects, or crowds of humans.

In this exercise we consider the motion of a single active particle in a viscous environment in two dimensions. To be specific, we consider a polar point particle at position $\vec{r}(t)$ that points into a direction given by a unit vector $\hat{u}(t) = (\cos\varphi(t), \sin\varphi(t))$ where the angle $\varphi(t)$ is given with respect to some arbitrary reference direction. The particle is subject to Brownian motion due to its surrounding. The activity is given by a force that acts along the direction of the particle, i.e., $\vec{F}_a(t) = f_a\hat{u}(t)$ with a constant amplitude f_a.

Due to the temperature, thermal forces $\vec{F}_T(t)$ and torques $T_T(t)$ act on the particle (details are specified later). Note that only rotations in the plane are considered, i.e., if considered as vectors all quantities related to rotations are perpendicular to the plane of motion.

Any motion of the particle is retarded by a friction force and torque given by Stokes' law and can be assumed to be $\vec{F}_S(t) = -\gamma\dot{\vec{r}}(t)$ and $T_S = -\gamma_R\dot{\varphi}(t)$ with friction constants γ and γ_R.

The mass of the particle is m and the moment of inertia is I.

a) [*1 point*] Write down the equations of motion for the position $\vec{r}(t)$ and angle $\varphi(t)$ given the forces and torques specified above.

b) [*2 points*] Without thermal forces or torques (this includes $T_S(0) = -\gamma_R\dot{\varphi}(0) = 0$), calculate the velocity $\dot{\vec{r}}(t)$ of the particle if the initial velocity is $\dot{\vec{r}}(t=0) = v_0\hat{u}_0$ and the initial direction $\hat{u}(t=0) = \hat{u}_0$. Explain your result phenomenologically.

c) [*1 point*] Determine conditions for the time t depending on the constants mentioned above such that the equations of motions from Task a) can be approximated by

$$\gamma\dot{\vec{r}}(t) = f_a\hat{u}(t) + \vec{F}_T(t),$$
$$\gamma_R\dot{\varphi}(t) = T_T(t).$$

Motivate your choice for the conditions. A strict calculation is not required. This is called the overdamped limit and will be used for the following tasks.

d) [*2 points*] Consider the overdamped equations given in Task (c). The thermal force and torque are considered to be random (corresponding to random Brownian kicks from the surrounding). If averaged over many runs one finds

$$\left\langle \vec{F}_T(t) \right\rangle = \vec{0},$$
$$\langle T_T(t) \rangle = 0,$$
$$\left\langle F_{T,j}(t) F_{T,k}(t') \right\rangle = 2\gamma k_B T \delta_{jk} \delta(t - t'),$$
$$\left\langle T_T(t) T_T(t') \right\rangle = 2\gamma_R k_B T \delta(t - t'),$$

where $k_B T$ corresponds to the thermal energy of the surrounding and j, k in the third line indicate the spatial component of the thermal forces.
Calculate $\langle \varphi(t) \rangle$ and $\left\langle \varphi^2(t) \right\rangle$ in case $\varphi(t = 0) = 0$ and $\dot{\varphi}(t = 0) = 0$.

e) [*4 points*] Consider that the system at $t = 0$ has relaxed into a state, where it is properly described by the overdamped equations given in Task (c). Calculate the mean positions $\langle \vec{r}(t) \rangle$ and mean-squared displacement $\left\langle |\vec{r}(t)|^2 \right\rangle$ for small times $t \ll 1$, i.e., up until terms with t^2, in case $\vec{r}(t = 0) = \vec{0}$, $\dot{\vec{r}}(t = 0) = f_a \hat{u}_0 / \gamma$, $\varphi(t = 0) = 0$, and $\dot{\varphi}(t = 0) = 0$.
Discuss the typical distance that a particle travels before it changes its direction significantly. For that consider first the time t_{reverse} it takes before the velocity changes from the initial condition $\dot{\vec{r}}(t = 0) = f_a \hat{u}_0 / \gamma$ to $\dot{\vec{r}}(t = t_{\text{reverse}}) = -f_a \hat{u}_0 / \gamma$. Note that the time scale related to this typical distance can be seen as the limit of validity of the small-time approximation used here.

Hint: *First calculate $\langle \hat{u}(t) \rangle$ (up to first order in t required) and $\langle \hat{u}(t) \hat{u}(t') \rangle$ (only zeroth order in t required).*

Solution

a) Newton's equations:

$$m\ddot{\vec{r}}(t) = -\gamma\dot{\vec{r}}(t) + f_a\hat{u}(t) + \vec{F}_T(t), \quad \text{[0.5 Points]}$$
$$I\ddot{\varphi}(t) = -\gamma_R\dot{\varphi}(t) + T_T(t). \quad \text{[0.5 Points]}$$

b) Without thermal torques the vector $\hat{u}(t)$ remains constant, i.e., $\hat{u}(t) = \hat{u}_0$. The equation $m\ddot{\vec{r}}(t) = -\gamma\dot{\vec{r}}(t) + f_a\hat{u}_0$ is solved by

$$\dot{\vec{r}}(t) = \vec{A}\exp(-t\gamma/m) + f_a\hat{u}_0/\gamma. \quad \text{[0.5 Points]}$$

The initial condition is fulfilled for $\vec{A} = (v_0 - f_a/\gamma)\,\hat{u}_0 \quad \text{[0.5 Points]}$.
Therefore, in the long-time limit the velocity $f_a\hat{u}_0/\gamma$ is approached [0.5 Points]. In case the initial velocity differs from this final velocity, the particle is accelerated or decelerated until the final velocity is reached [0.5 Points]. *Additional comments: The approach to the final velocity is given by an exponential decay and the time-scale of the approach is given by* m/γ.

c) The overdamped limit is valid in the long-time limit, i.e., for $t > m/\gamma$ and $t > I/\gamma_R$ [0.5 Points]. For this choice the inertial terms can be neglected as they are smaller than the friction terms [0.5 Points].

d) From $\gamma_R\dot{\varphi}(t) = T_T(t)$ one finds $\varphi(t) = \int_0^t dt' T_T(t')/\gamma_R$ and therefore $\langle\varphi(t)\rangle = \int_0^t dt' \langle T_T(t')\rangle/\gamma_R = 0$ [0.5 Points]. Furthermore,

$$\begin{aligned}
\left\langle (\varphi(t))^2 \right\rangle &= \frac{1}{\gamma_R^2}\left\langle \int_0^t dt' T_T(t') \int_0^t dt'' T_T(t'') \right\rangle \\
&= \frac{1}{\gamma_R^2}\int_0^t dt' \int_0^t dt'' \langle T_T(t')T_T(t'')\rangle \quad \text{[0.5 Points]} \\
&= \frac{1}{\gamma_R^2}\int_0^t dt' \int_0^t dt'' 2\gamma_R k_B T\delta(t'-t'') \\
&= \frac{1}{\gamma_R^2}\int_0^t dt' 2\gamma_R k_B T \quad \text{[0.5 Points]} \\
&= \frac{2k_BT}{\gamma_R}t. \quad \text{[0.5 Points]}
\end{aligned}$$

e)

$$\begin{aligned}
\langle\hat{u}(t)\rangle &= \langle(\cos\varphi(t), \sin\varphi(t))\rangle \\
&\approx \left\langle \left(1 - \varphi^2(t)/2, \varphi(t)\right) \right\rangle \\
&= \left(1 - \left\langle\varphi^2(t)\right\rangle/2, \langle\varphi(t)\rangle\right) \\
&= \left(1 - \frac{k_BT}{\gamma_R}t, 0\right) \quad \text{[1 Points]}
\end{aligned}$$

Due to $\gamma\dot{\vec{r}}(t) = f_a\hat{u}(t) + \vec{F}_T(t)$ it is $\vec{r}(t) = \int_0^t dt' \left(f_a\hat{u}(t') + \vec{F}_T(t')\right)/\gamma$ and therefore

$\langle \vec{r}(t) \rangle = f_a \int_0^t dt' \langle \hat{u}(t') \rangle / \gamma + 0$. As a consequence $\langle y(t) \rangle = 0$ and

$$\begin{aligned} \langle x(t) \rangle &= f_a \int_0^t dt' \left(1 - \frac{k_B T}{\gamma_R} t \right) / \gamma \\ &= \frac{f_a}{\gamma} t - \frac{f_a k_B T}{2 \gamma \gamma_R} t^2. \quad \text{[0.5 Points]} \end{aligned}$$

Therefore, the typical time to change the direction is $\frac{\gamma_R}{k_B T}$. As the particle travels with a mean velocity $\frac{f_a}{\gamma}$ the typical length (persistence length) is $\frac{f_a \gamma_R}{\gamma k_B T}$ [0.5 Points]. (*Side note: You might define the persistence length with some additional constant prefactor like* $\ln(2)$. *That is fine as well, as no further details are given in the exercise.*)

Furthermore,

$$\langle \hat{u}(t) \hat{u}(t') \rangle \approx 1 - \left\langle \left(\varphi(t) - \varphi(t') \right)^2 / 2 \right\rangle \approx 1. \quad \text{[0.5 Points]}$$

Due to $\gamma \dot{\vec{r}}(t) = f_a \hat{u}(t) + \vec{F}_T(t)$ it is $\vec{r}(t) = \int_0^t dt' \left(f_a \hat{u}(t') + \vec{F}_T(t') \right) / \gamma$ and therefore (the terms odd in $\hat{u}(t)$ and $\vec{F}_T(t')$ are left away as it is 0 after taking the average)

$$\begin{aligned} \left\langle |\vec{r}(t)|^2 \right\rangle &= \frac{f_a^2}{\gamma^2} \int_0^t dt' \int_0^t dt'' \langle \hat{u}(t') \hat{u}(t'') \rangle + \frac{1}{\gamma^2} \int_0^t dt' \int_0^t dt'' \left\langle \vec{F}_T(t') \vec{F}_T(t'') \right\rangle \quad \text{[0.5 Points]} \\ &= \frac{f_a^2}{\gamma^2} \int_0^t dt' \int_0^t dt'' 1 + \frac{4 k_B T}{\gamma} \int_0^t dt' \int_0^t dt'' \delta(t' - t'') \quad \text{[0.5 Points]} \\ &= \frac{f_a^2}{\gamma^2} t^2 + \frac{4 k_B T}{\gamma} t \quad \text{[0.5 Points]} \end{aligned}$$

Further corrections are at least of order t^3.

Problem 8

Quantum Convolutional Neural Network

Dr. Petr Zapletal, Timo Eckstein, and Prof. Dr. Michael J. Hartmann – *Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)*

Background Convolutional neural networks are artificial neural networks, which can be exploited for image recognition. They can process rasterized images and classify whether, for example, a *cat* or a *dog* is depicted in a particular image. Quantum convolutional neural networks (QCNNs) are analogous to their classical counterparts and they are designed to recognize quantum phases of matter, which are characterized by long-range quantum correlations. QCNNs process quantum states in order to determine whether they belong to a given quantum phase. Here we consider a minimal example of a QCNN, which processes a quantum state of four qubits. The Hilbert space of each qubit is spanned by two quantum states $|0\rangle_i$ and $|1\rangle_i$ for $i = 1, 2, 3, 4$. The QCNN is based on a quantum circuit depicted in Fig. 8.1, which consists of a convolutional (C) layer, a pooling (P) layer and the measurement of qubits 1 and 4 at the end of the circuit. In the C layer, a translationally invariant unitary transformation U_C consisting of CZ gates between neighboring qubits is performed. In the P layer, a unitary transformation U_P consisting of two CNOT gates is performed and qubit 2 as well as 3 are discarded such that only qubits 1 and 4 are measured at the end of the circuit.

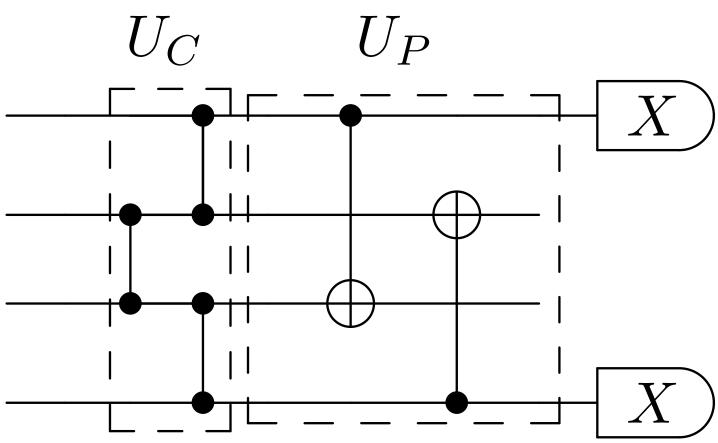


Figure 8.1 Quantum convolutional neural network consisting of the convolutional layer U_C, the pooling layer U_P and the measurement of qubits 1 and 4. Horizontal lines represent qubits, which are evolved in the circuit from left to right. CZ gates are depicted as vertical lines acting on two qubits denoted by dots. CNOT gates are depicted as vertical lines with a control qubit denoted by a dot and a target qubit denoted by a cross.

A CNOT$_{ij}$ gate with a control qubit i and a target qubit j performs a bit flip $|0\rangle_j \leftrightarrow |1\rangle_j$ on the target qubit if the control qubit is in the state $|1\rangle_i$ and it does not perform any transformation if the control qubit is in the state $|0\rangle_i$. A CZ$_{ij}$ gate acting on qubits i and j induces a phase shift $|11\rangle_{ij} \to -|11\rangle_{ij}$ if the two qubits are in the state $|11\rangle_{ij}$ and it does not perform any transformation otherwise, where we use the notation $|\psi\theta\rangle_{ij} = |\psi\rangle_i \otimes |\theta\rangle_j$ for the tensor product of two states.

The QCNN is designed to recognize the $(\mathbb{Z}_2 \times \mathbb{Z}_2)$ symmetry-protected topological (SPT) phase. The states belonging to this quantum phase are characterized by a non-vanishing expectation value of so-called *string order parameters*. For the four-qubit system considered here, the string order parameters are $\mathcal{O}_1 = X_1X_3Z_4$ and $\mathcal{O}_2 = Z_1X_2X_4$. For an input state $|\psi\rangle$, the QCNN output $x_{\text{QCNN}} = \langle\psi| U^\dagger \frac{X_1+X_4}{2} U |\psi\rangle$ is the expectation value of $(X_1 + X_4)/2$ measured after the QCNN circuit $U = U_P U_C$. Hence, the QCNN output corresponds to the expectation value $\langle\psi| \mathcal{O} |\psi\rangle$ of the observable $\mathcal{O} = U^\dagger \frac{X_1+X_4}{2} U$ measured directly on the input state $|\psi\rangle$.

a) [*2 points*] Prove the following gate identities

$$Z_i^\dagger X_i Z_i = -X_i, \tag{8.1}$$
$$[\text{CZ}_{ij}, Z_i] = [\text{CZ}_{ij}, Z_j] = 0, \tag{8.2}$$
$$\text{CZ}_{ij}^\dagger X_i \text{CZ}_{ij} = X_i Z_j, \tag{8.3}$$
$$\text{CZ}_{ij}^\dagger X_j \text{CZ}_{ij} = Z_i X_j, \tag{8.4}$$
$$\text{CNOT}_{ij}^\dagger X_i \text{CNOT}_{ij} = X_i X_j, \tag{8.5}$$

where X_i and Z_i are Pauli operators with eigenstates $|\pm\rangle_i = (|0\rangle_i \pm |1\rangle_i)/\sqrt{2}$ and $|0/1\rangle_i$, respectively, and $[A, B] = AB - BA$ is the commutator.
Hint: use the matrix representation of the quantum states

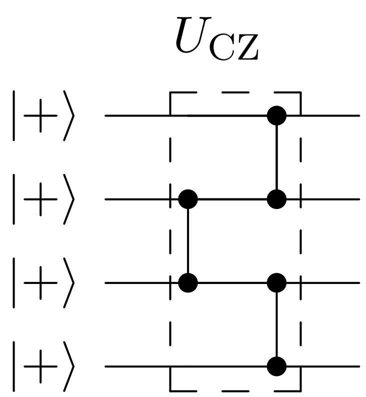
$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow |00\rangle, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow |01\rangle, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \leftrightarrow |10\rangle, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \leftrightarrow |11\rangle, \tag{8.6}$$

and Pauli operators

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \leftrightarrow X, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \leftrightarrow Z. \tag{8.7}$$

b) [*1 point*] Use identities proven in part **1)** to express the observable $\mathcal{O}$ in terms of the string order parameters $\mathcal{O}_1$ and $\mathcal{O}_2$ defined above.

c) [*1 point*] We now consider the so-called *cluster state* $|C\rangle$ which is uniquely defined as an eigenstate of so-called *stabilizer generators* $S_1 = X_1Z_2$, $S_2 = Z_1X_2Z_3$, $S_3 = Z_2X_3Z_4$, and $S_4 = Z_3X_4$ with the eigenvalue $+1$ for all four stabilizer generators. Show that the circuit U_{CZ}



consisting of CZ gates between neighboring qubits, prepares the cluster state from the product state $|++++\rangle = |+\rangle_1 \otimes |+\rangle_2 \otimes |+\rangle_3 \otimes |+\rangle_4$.

d) [*1 point*] The QCNN output x_{QCNN} is equal to unity for states belonging to the SPT phase and it is equal to zero for states that do not belong to the SPT phase. Using the equivalence $x_{\mathrm{QCNN}} = \langle \mathcal{O} \rangle$, determine whether the cluster state $|C\rangle$ and states $|P\rangle = |++++\rangle$ as well as $|A\rangle = |+-+-\rangle$ belong to the SPT phase.

e) [*2 points*] An important feature of the QCNN is that it can tolerate certain type of perturbations $\mathcal{P}$ such that a perturbed state $\left|\tilde{\psi}\right\rangle = \mathcal{P}\,|\psi\rangle$ retains the same QCNN output x_{QCNN} as an unperturbed state $|\psi\rangle$ for any $|\psi\rangle$. Determine, which of the single qubit perturbations

$$\mathcal{P} \in \{X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4, Z_1, Z_2, Z_3, Z_4\} \tag{8.8}$$

are tolerated by the QCNN, where the matrix representation of the Pauli Y_j operator is

$$Y_j \leftrightarrow \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{8.9}$$

f) [*3 points*] Find all states belonging to the SPT phase that yield the QCNN output $x_{QCNN} = 1$.

Solution

a) We prove the gate identities using the matrix representation

$$Z_i^\dagger X_i Z_i \leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = (-1)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \leftrightarrow -X, \tag{8.10}$$

$$[\mathrm{CZ}_{ij}, Z_i] \leftrightarrow \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1 \end{bmatrix} - \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix} = 0 \tag{8.11}$$

$$[\mathrm{CZ}_{ij}, Z_j] \leftrightarrow \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix} - \begin{bmatrix} 1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix} = 0 \tag{8.12}$$

$$\mathrm{CZ}_{ij}^\dagger X_i \mathrm{CZ}_{ij} \leftrightarrow \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix} = \begin{bmatrix} 0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0 \end{bmatrix} \leftrightarrow X_i Z_j, \tag{8.13}$$

$$\mathrm{CZ}_{ij}^\dagger X_j \mathrm{CZ}_{ij} \leftrightarrow \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix}\begin{bmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1 \end{bmatrix} = \begin{bmatrix} 0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0 \end{bmatrix} \leftrightarrow Z_i X_j, \tag{8.14}$$

$$\mathrm{CNOT}_{ij}^\dagger X_i \mathrm{CNOT}_{ij} \leftrightarrow \begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0 \end{bmatrix}\begin{bmatrix} 0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0 \end{bmatrix}\begin{bmatrix} 1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0 \end{bmatrix} = \begin{bmatrix} 0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0 \end{bmatrix} \leftrightarrow X_i X_j. \tag{8.15}$$

Alternatively, one can prove Eq. (8.1) by using that Z_i and X_i anti-commute (the anti-commutation does not need to be proven) and showing

$$Z_i^\dagger X_i Z_i = -Z_i^\dagger Z_i X_i = -X_i. \tag{8.16}$$

To prove Eq. (8.2), one can alternatively argue that Z_i, Z_j and CZ_{ij} are all diagonal in the computational basis and thus mutually commute with each other.

2 points in total: half a point for proving Eq. 8.1; half a point for proving Eq. 8.2; half a point for proving Eqs. 8.3 and (8.4); half a point for proving Eq. 8.5.

b) We use gate identities (8.3), (8.4), and (8.5) and that $Z_i^2 = \mathbb{1}$ to show that

$$\mathcal{O} = U_C^\dagger U_P^\dagger \frac{X_1 + X_4}{2} U_P U_C = \frac{1}{2}\left(U_C^\dagger X_1 X_3 U_C + U_C^\dagger X_2 X_4 U_C\right) \tag{8.17}$$

$$= \frac{1}{2}\left(X_1 Z_2^2 X_3 Z_4 + Z_1 X_2 Z_3^2 X_4\right) = \frac{1}{2}\left(X_1 X_3 Z_4 + Z_1 X_2 X_4\right) = \frac{1}{2}\left(\mathcal{O}_1 + \mathcal{O}_2\right). \tag{8.18}$$

1 point in total: half a point for correctly using the identities (8.3), (8.4); another half a point for correctly using the identity (8.5) and $Z_i^2 = \mathbb{1}$.

c) We use gate identities (8.2), (8.3), and (8.4), that $U_{\text{CZ}} U_{\text{CZ}}^\dagger = \mathbb{1}$, that $Z_i^2 = \mathbb{1}$ and that $X_i \left|+\right\rangle_i = \left|+\right\rangle_i$ to show that

$$S_1 \left|C\right\rangle = U_{\text{CZ}} U_{\text{CZ}}^\dagger X_1 Z_2 U_{\text{CZ}} \left|++++\right\rangle = U_{\text{CZ}} X_1 Z_2^2 \left|++++\right\rangle = U_{\text{CZ}} \left|++++\right\rangle = \left|C\right\rangle, \tag{8.19}$$

$$S_2 \left|C\right\rangle = U_{\text{CZ}} U_{\text{CZ}}^\dagger Z_1 X_2 Z_3 U_{\text{CZ}} \left|++++\right\rangle = U_{\text{CZ}} Z_1^2 X_2 Z_3^2 \left|++++\right\rangle = U_{\text{CZ}} \left|++++\right\rangle = \left|C\right\rangle, \tag{8.20}$$

$$S_3 \left|C\right\rangle = U_{\text{CZ}} U_{\text{CZ}}^\dagger Z_2 X_3 Z_4 U_{\text{CZ}} \left|++++\right\rangle = U_{\text{CZ}} Z_2^2 X_3 Z_4^2 \left|++++\right\rangle = U_{\text{CZ}} \left|++++\right\rangle = \left|C\right\rangle, \tag{8.21}$$

$$S_4 \left|C\right\rangle = U_{\text{CZ}} U_{\text{CZ}}^\dagger Z_3 X_4 U_{\text{CZ}} \left|++++\right\rangle = U_{\text{CZ}} Z_3^2 X_4 \left|++++\right\rangle = U_{\text{CZ}} \left|++++\right\rangle = \left|C\right\rangle. \tag{8.22}$$

We conclude that the state $U_{\text{CZ}} \left|++++\right\rangle$ prepared from the product state $\left|++++\right\rangle$ by the circuit U_{CZ} is an eigenstate with the eigenvalue $+1$ of all four stabilizer generators and it is thus the cluster state.

1 point in total for showing that the cluster state is an eigenstate with the eigenvalue $+1$ of all four stabilizer generators: half a point for transforming the stabilizer generators according to U_{CZ}; another half a point for using $Z_i^2 = \mathbb{1}$ and that $X_i \left|+\right\rangle_i = \left|+\right\rangle_i$ to complete the calculations.

d) We start by using that the cluster state is the $+1$ eigenstate of the four stabilizer generators to show that

$$\mathcal{O}_1 \left|C\right\rangle = X_1 Z_2 Z_2 X_3 Z_4 \left|C\right\rangle = S_1 S_3 \left|C\right\rangle = \left|C\right\rangle, \tag{8.23}$$

$$\mathcal{O}_2 \left|C\right\rangle = Z_1 X_2 Z_3 Z_3 X_4 \left|C\right\rangle = S_2 S_4 \left|C\right\rangle = \left|C\right\rangle, \tag{8.24}$$

from which it follows that

$$x_{QCNN} = \left\langle C\right| \frac{\mathcal{O}_1 + \mathcal{O}_2}{2} \left|C\right\rangle = \left\langle C | C\right\rangle = 1, \tag{8.25}$$

and that the cluster state belongs to the SPT phase. For the other two states, we obtain

$$x_{QCNN} = \left\langle ++++\right| \frac{\mathcal{O}_1 + \mathcal{O}_2}{2} \left|++++\right\rangle = \frac{1}{2}\left(\left\langle ++++|+++-\right\rangle + \left\langle ++++|-+++\right\rangle\right) = 0, \tag{8.26}$$

$$x_{QCNN} = \left\langle +-+-\right| \frac{\mathcal{O}_1 + \mathcal{O}_2}{2} \left|+-+-\right\rangle = \frac{1}{2}\left(\left\langle +-+-|+-++\right\rangle + \left\langle +-+-|--+-\right\rangle\right) = 0, \tag{8.27}$$

showing that they do not belong to the SPT phase, where we used that $X_i \left|+\right\rangle_i = \left|+\right\rangle_i$ and $Z_i \left|+\right\rangle_i = \left|-\right\rangle_i$.

1 point in total: half a point for showing that the cluster state belongs to the SPT phase; another half a point for showing that the other two states do not belong to the SPT phase.

e) For a given perturbation $\mathcal{P}$, the QCNN outputs $\langle\psi|\,\mathcal{O}\,|\psi\rangle$ and $\langle\psi|\,\mathcal{P}^\dagger\mathcal{O}\mathcal{P}\,|\psi\rangle$ for a perturbed state $\mathcal{P}\,|\psi\rangle$ and an unperturbed state $|\psi\rangle$, respectively, coincide for any $|\psi\rangle$ *if and only if* $\mathcal{P}^\dagger\mathcal{O}\mathcal{P} = \mathcal{O}$, which is equivalent to $[\mathcal{O},\mathcal{P}] = 0$. We start by recalling that $[\sigma_i^\alpha, \sigma_j^\beta] = 0$ for $i \neq j$ as well as any $\alpha = X, Y, Z$ and $\beta = X, Y, Z$ and that $\sigma_i^\alpha\sigma_i^\beta = -\sigma_i^\beta\sigma_i^\alpha$ for any $\alpha \neq \beta$, where $\sigma_i^X = X_i$, $\sigma_i^Y = Y_i$, and $\sigma_i^Z = Z_i$. The string order parameters $\mathcal{O}_1$ and $\mathcal{O}_2$ either commute or anti-commute with the single qubit perturbations $\mathcal{P}$. For example

$$X_1\mathcal{O}_1 = X_1X_1X_3Z_4 = X_1X_3Z_4X_1 = \mathcal{O}_1X_1, \tag{8.28}$$
$$X_1\mathcal{O}_2 = X_1Z_1X_2X_4 = -Z_1X_2X_4X_1 = -\mathcal{O}_2X_1. \tag{8.29}$$

As $\mathcal{O}$ is a sum of the string order parameters $\mathcal{O}_1$ and $\mathcal{O}_2$, only perturbations commuting with both $\mathcal{O}_1$ and $\mathcal{O}_2$ are tolerated by the QCNN. We thus conclude that only perturbations $\mathcal{P} = X_2$ and $\mathcal{P} = X_3$ are tolerated by the QCNN. All other single qubit perturbations modify the QCNN output

$$X_1^\dagger\mathcal{O}X_1 = \frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.30}$$
$$X_4^\dagger\mathcal{O}X_4 = -\frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.31}$$
$$Y_1^\dagger\mathcal{O}Y_1 = -\frac{\mathcal{O}_1 + \mathcal{O}_2}{2}, \tag{8.32}$$
$$Y_2^\dagger\mathcal{O}Y_2 = \frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.33}$$
$$Y_3^\dagger\mathcal{O}Y_3 = -\frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.34}$$
$$Y_4^\dagger\mathcal{O}Y_4 = -\frac{\mathcal{O}_1 + \mathcal{O}_2}{2}, \tag{8.35}$$
$$Z_1^\dagger\mathcal{O}Z_1 = -\frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.36}$$
$$Z_2^\dagger\mathcal{O}Z_2 = \frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.37}$$
$$Z_3^\dagger\mathcal{O}Z_3 = -\frac{\mathcal{O}_1 - \mathcal{O}_2}{2}, \tag{8.38}$$
$$Z_4^\dagger\mathcal{O}Z_4 = \frac{\mathcal{O}_1 - \mathcal{O}_2}{2}. \tag{8.39}$$

where we used the commutation relations stated above and that $(\sigma_i^\alpha)^2 = \mathbb{1}$. For example, the perturbed cluster state $Z_{2/3}\,|C\rangle$ yields a vanishing QCNN output

$$x_{\text{QCNN}} = \langle C|\,Z_{2/3}^\dagger\mathcal{O}Z_{2/3}\,|C\rangle = 0. \tag{8.40}$$

2 points in total: half a point for realizing that tolerated perturbations commute with $\mathcal{O}$; half a point for realizing that $\mathcal{O}_1$ and $\mathcal{O}_2$ either commute or anti-commute with all perturbations; one point for correctly identifying tolerated perturbations.

f) In order to yield the QCNN output $x_{\text{QCNN}} = 1$, a state needs to be a $+1$ eigenstate of both $\mathcal{O}_1$ and $\mathcal{O}_2$. Using identities (8.2), (8.3) and (8.4) as well as that $Z_i^2 = \mathbb{1}$, we show that the unitary U_{CZ} transforms the string order parameters $U_{\rm CZ}^\dagger \mathbb{O}_1 U_{\rm CZ} = X_1X_3$ and $U_{\rm CZ}^\dagger \mathbb{O}_2 U_{\rm CZ} = X_2X_4$ into diagonal operators in the X-basis. Inspecting X-basis eigenstates, we find that four states $|++++\rangle$, $|-+-+\rangle$, $|+-+-\rangle$, and $|----\rangle$ are eigenstates with the eigenvalue $+1$ for both diagonal operators X_1X_3 and X_2X_4. All other X-basis eigenstates have at least one eigenvalue -1. We thus obtain

$$\mathbb{O}_1 U_{\rm CZ}|++++\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_1 U_{\rm CZ}|++++\rangle = U_{\rm CZ}X_1X_3|++++\rangle = U_{\rm CZ}|++++\rangle, \tag{8.41}$$

$$\mathbb{O}_1 U_{\rm CZ}|-+-+\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_1 U_{\rm CZ}|-+-+\rangle = U_{\rm CZ}X_1X_3|-+-+\rangle = U_{\rm CZ}|-+-+\rangle, \tag{8.42}$$

$$\mathbb{O}_1 U_{\rm CZ}|+-+-\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_1 U_{\rm CZ}|+-+-\rangle = U_{\rm CZ}X_1X_3|+-+-\rangle = U_{\rm CZ}|+-+-\rangle, \tag{8.43}$$

$$\mathbb{O}_1 U_{\rm CZ}|----\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_1 U_{\rm CZ}|----\rangle = U_{\rm CZ}X_1X_3|----\rangle = U_{\rm CZ}|----\rangle, \tag{8.44}$$

$$\mathbb{O}_2 U_{\rm CZ}|++++\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_2 U_{\rm CZ}|++++\rangle = U_{\rm CZ}X_2X_4|++++\rangle = U_{\rm CZ}|++++\rangle, \tag{8.45}$$

$$\mathbb{O}_2 U_{\rm CZ}|-+-+\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_2 U_{\rm CZ}|-+-+\rangle = U_{\rm CZ}X_2X_4|-+-+\rangle = U_{\rm CZ}|-+-+\rangle, \tag{8.46}$$

$$\mathbb{O}_2 U_{\rm CZ}|+-+-\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_2 U_{\rm CZ}|+-+-\rangle = U_{\rm CZ}X_2X_4|+-+-\rangle = U_{\rm CZ}|+-+-\rangle, \tag{8.47}$$

$$\mathbb{O}_2 U_{\rm CZ}|----\rangle = U_{\rm CZ}U_{\rm CZ}^\dagger \mathbb{O}_2 U_{\rm CZ}|----\rangle = U_{\rm CZ}X_2X_4|----\rangle = U_{\rm CZ}|----\rangle, \tag{8.48}$$

showing that $|C_1\rangle = |C\rangle = U_{\rm CZ}|++++\rangle$, $|C_2\rangle = U_{\rm CZ}|-+-+\rangle$, $|C_3\rangle = U_{\rm CZ}|+-+-\rangle$, and $|C_4\rangle = U_{\rm CZ}|----\rangle$ are four eigenstates of the string order parameters with the eigenvalue $+1$ for both string order parameters. Thanks to the mapping onto operators X_1X_3 and X_2X_4, we know that there are no other $+1$ eigenstates of both string order parameters and that the eigenstates are orthogonal $\langle C_i|C_k\rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta. We thus conclude that states $|C_i\rangle$, for $i = 1,2,3,4$, form a basis of the subspace spanned by all states belonging to the SPT phase yielding $x_{\rm QCNN} = 1$. That means that any state $|\psi\rangle$ with $x_{\rm QCNN} = 1$ can be expressed as a superposition $|\psi\rangle = \sum_{i=1}^4 \gamma_i |C_i\rangle$.

Alternative solution: we first realize that the string order parameters are parity operators $\mathbb{O}_1^2 = \mathbb{O}_2^2 = \mathbb{1}$. Each of the parity operators splits the Hilbert space into two subspaces of equal dimensions spanned by $+1$ and -1 eigenstates. We thus conclude that the subspace of their common $+1$ eigenstates has the dimension $d = 2^4/2^2 = 4$. Using the result of part **5)** that perturbations X_2 and X_3 do not modify the QCNN output, we deduce that states $|C_2\rangle = X_2|C\rangle$, $|C_3\rangle = X_3|C\rangle$, and $|C_4\rangle = X_2X_3|C\rangle$ yield the QCNN output $x_{\rm QCNN} = 1$. At the same time, these states are orthogonal to the cluster state

$$\langle C|C_2\rangle = \langle ++++|U_{\rm CZ}^\dagger X_2 U_{\rm CZ}|++++\rangle = \langle ++++|Z_1X_2Z_3|++++\rangle = 0 \tag{8.49}$$

$$\langle C|C_3\rangle = \langle ++++|U_{\rm CZ}^\dagger X_3 U_{\rm CZ}|++++\rangle = \langle ++++|Z_2X_3Z_4|++++\rangle = 0 \tag{8.50}$$

$$\langle C|C_4\rangle = \langle ++++|U_{\rm CZ}^\dagger X_2X_3 U_{\rm CZ}|++++\rangle = \langle ++++|Z_1X_2Z_2Z_3X_3Z_4|++++\rangle = 0. \tag{8.51}$$

Analogously, it can be shown that they are orthogonal to each other. We thus conclude that states $|C\rangle_1 = |C\rangle$, $|C\rangle_2$, $|C\rangle_3$, and $|C\rangle_4$ form a basis of the four-dimensional subspace of all states belonging to the SPT phase yielding $x_{\rm QCNN} = 1$.

3 points in total: half a point for realizing that $+1$ eigenstates of both $\mathbb{O}_1$ and $\mathbb{O}_2$ yield $x_{\rm QCNN} = 1$; one point for finding four states $|C_i\rangle$ yielding $x_{\rm QCNN} = 1$; half a point for showing that they are orthogonal (either by the mapping onto X-basis eigenstates or by a direct inspection); one point for arguing that any state $|\psi\rangle$ with $x_{\rm QCNN} = 1$ can be expressed as a superposition $|\psi\rangle = \sum_{i=1}^4 \gamma_i |C_i\rangle$ (either by mapping $\mathbb{O}_1$ and $\mathbb{O}_2$ onto diagonal operators and inspecting their eigenstates or by identifying the dimension of the common $+1$ subspace of parity operators $\mathbb{O}_1$ and $\mathbb{O}_2$).

Problem 9

Hawking Radiation, the Logarithmic Phase Singularity, and the Inverted Harmonic Oscillator

Freyja Ullinger[1,2], Dr. Matthias Zimmermann[2], and Prof. Dr. Wolfgang P. Schleich[1] – [1]*Institut für Quantenphysik, Ulm University,* [2]*Institute of Quantum Technologies, German Aerospace Center (DLR)*

Background A spacetime singularity is located at the center of a black hole and surrounded by an event horizon, separating spacetime into two disjunct regions: one of them accessible to an outside observer and one that is not. At the event horizon, a logarithmic phase singularity emerges in the mode functions of a massless scalar field, being characteristic for Hawking radiation emitted by the black hole.

There are many situations when physical systems display phenomena connected to black hole evaporation. They range from acceleration radiation over the presence of a sonic horizon for sound waves, the quantum catastrophe of slow light and Bose-Einstein condensates, to setups employing water waves. Insight into this plethora of physical systems can be provided by simple models that cover the main features of the underlying effects. In view of Hawking radiation such an elementary model is the inverted harmonic oscillator.

Overview

In the first part of this exercise, we consider a *classical* particle of mass m exposed to an *inverted harmonic oscillator* of steepness $\omega > 0$, as described by the potential

$$V(x) = -\frac{1}{2}m\omega^2 x^2 \tag{9.1}$$

which depends on the coordinate x and is depicted in Fig. 9.1. You will show that, similar to the event horizons of a black hole, also an inverted harmonic oscillator displays horizons, which are however located in phase space instead of spacetime.

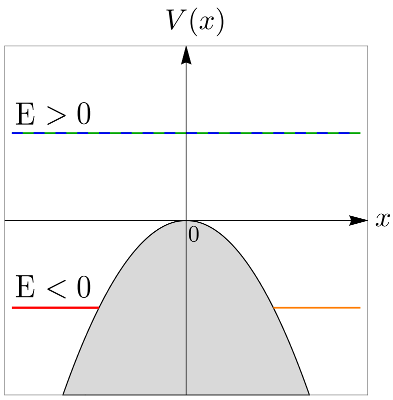


Figure 9.1 The inverted harmonic oscillator potential $V(x)$, Eq. (9.1), as a function of the position x. For each energy E we depict two cases corresponding to an incoming classical particle from the left or right, respectively. A classical particle with negative energy $E < 0$ (red and orange lines) is reflected at the potential barrier. On the contrary, a particle with positive energy $E > 0$ (blue and green lines) is able to surpass it. Figure reprinted from F. Ullinger et al., AVS Quantum Sci. **4**, 024402 (2022) under license CC BY 4.0.

Next, we turn to the *quantum* system of an inverted oscillator and analyze its properties. In particular, you will reveal a *logarithmic phase singularity* in this system and demonstrate that it causes a transmission and reflection coefficient which resembles a very particular quantum statistics.

Afterwards, we identify a logarithmic phase singularity emerging in the mode functions of an electromagnetic field at the event horizon of a black hole. You will show that this particular singularity is a characteristic feature of Hawking radiation.

As a result of your efforts, you will get a glimpse into the intriguing similarities between Hawking radiation emitted at the event horizon of a black hole and the simple system of an inverted harmonic oscillator.

a) [*1.5 points*] First, we consider a one-dimensional inverted harmonic oscillator with steepness ω as characterized by the potential $V(x)$, Eq. (9.1), displayed in Fig. 9.1. In this system, the dynamics of a *classical particle* of mass m is governed by the Hamiltonian

$$H(x,p) = \frac{p^2}{2m} - \frac{1}{2}m\omega^2 x^2 \tag{9.2}$$

with position x and momentum p. Since the Hamiltonian $H(x,p)$, Eq. (9.2), is time-independent, each classical trajectory is associated with a particular energy $E = H(x_0, p_0)$ as determined by the initial conditions $x(0) = x_0$ and $p(0) = p_0$ for the respective motion at time $t = 0$.

1. Identify in a sketch of phase space, *i.e.* the two-dimensional space of position x and momentum p, the regions with phase space trajectories of energy (i) $E < 0$, (ii) $E = 0$, and (iii) $E > 0$.
2. Sketch a phase space trajectory $\{x(t), p(t)\}$ in each quadrant of phase space. How many distinct trajectories exist for a given energy E?

Hint: *In the first task, the relevant quadrants of phase space are identified as regions associated with different energy domains.*

b) [*1.5 points*] In order to make contact with phenomena familiar from black hole evaporation, we introduce the horizon coordinates

$$\xi \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{p}{m\omega}\right) \tag{9.3}$$

and

$$\eta \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{p}{m\omega}\right), \tag{9.4}$$

where $\hbar$ denotes the reduced Planck constant. We label the coordinates $\xi = 0$ and $\eta = 0$ as the *horizons in phase space*. In this particular basis the Hamiltonian H, Eq. (9.2), takes the form

$$H = -\frac{\hbar\omega}{2}\left(\xi\eta + \eta\xi\right). \tag{9.5}$$

1. Consider a classical particle of mass m in the inverted harmonic oscillator that is initially located at the coordinate $\xi_0 = \xi(0)$ at time $t = 0$. Determine the time T_1 which the particle requires to arrive at the coordinate $\xi_1 = \xi(T_1)$ as a function of the initial and final coordinate. For this purpose, solve the equations of motion

$$\dot{\xi} = \frac{1}{\hbar}\frac{\partial H}{\partial \eta}, \tag{9.6}$$

$$\dot{\eta} = -\frac{1}{\hbar}\frac{\partial H}{\partial \xi}. \tag{9.7}$$

2. How long does it take for the particle to reach the horizon $\xi = 0$ if $\xi_0 \neq 0$? How does your result depend on the value of $\eta_0 = \eta(0)$ at time $t = 0$?

3. What happens to the horizons in phase space in the limit $\omega \to 0$? What does this tell you about the allowed values of the energy E in this limit? Support your statements with the help of a phase space sketch.

c) [*0.5 points*] In the following, we consider a *quantum particle* of mass m subject to an inverted harmonic oscillator potential $V(x)$, Eq. (9.1). For our analysis, we make use of the operators

$$\hat{\xi} \equiv \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{\hat{p}}{m\omega} \right) \tag{9.8}$$

and

$$\hat{\eta} \equiv \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{\hat{p}}{m\omega} \right), \tag{9.9}$$

defined in terms of the position operator $\hat{x}$ and the momentum operator $\hat{p}$.

Determine the commutator $\left[\hat{\xi}, \hat{\eta}\right]$, provided that the position operator $\hat{x}$ and the momentum operator $\hat{p}$ satisfy the standard commutation relation. What does this tell you about the relationship between the operators $\hat{\xi}$ and $\hat{\eta}$?

d) [*1 point*] In analogy to the classical situation, see Eq. (9.5), the quantum mechanical Hamiltonian for an inverted harmonic oscillator takes the form

$$\hat{H} = -\frac{\hbar\omega}{2} \left(\hat{\xi}\hat{\eta} + \hat{\eta}\hat{\xi} \right). \tag{9.10}$$

We are now interested in the energy eigenstates $|\varepsilon\rangle$ of this system which are solutions of the stationary Schrödinger equation

$$\hat{H} |\varepsilon\rangle = \hbar\omega\varepsilon |\varepsilon\rangle \tag{9.11}$$

with real-valued dimensionless energy ε.

By making use of the ξ-representation, show for $\xi \neq 0$ that the two degenerate wave functions

$$\Psi_\varepsilon^\pm(\xi) = \frac{1}{\sqrt{2\pi\,|\xi|}} \exp\left(-\mathrm{i}\varepsilon \ln|\xi|\right) \Theta(\pm\xi) \tag{9.12}$$

are solutions of Eq. (9.11), which are governed by a *logarithmic phase singularity* at the horizon $\xi = 0$ in phase space. Here we have introduced the Heaviside step function

$$\Theta(x) \equiv \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{9.13}$$

Hint: *Analogous to the position representation, the ξ-representation $\Psi_\varepsilon^\pm(\xi) = \langle\xi|\Psi_\varepsilon^\pm\rangle$ of a quantum state is defined by making use of the eigenstates $|\xi\rangle$ of the operator $\hat{\xi}$, Eq. (9.8), satisfying the eigenvalue equation $\hat{\xi}\,|\xi\rangle = \xi\,|\xi\rangle$.*

e) [*0.5 points*] By solving Eq. (9.11), determine the degenerate eigenstates $|\Phi_\varepsilon^\pm\rangle$ with dimensionless energy ε, whose η-representation $\langle\eta|\Phi_\varepsilon^\pm\rangle$ is proportional to the Heaviside step function $\Theta(\pm\eta)$.

f) [*1.5 points*] Show that the quantum state $|\Psi_\varepsilon^+\rangle$, see Eq. (9.12), can be expressed as a superposition

$$\left|\Psi_\varepsilon^+\right\rangle = S_+(\varepsilon)\left|\Phi_\varepsilon^+\right\rangle + S_-(\varepsilon)\left|\Phi_\varepsilon^-\right\rangle \tag{9.14}$$

of the states $|\Phi_\varepsilon^\pm\rangle$ with energy-dependent coefficients

$$S_\pm(\varepsilon) = \frac{\Gamma\left(\frac{1}{2} - \mathrm{i}\varepsilon\right)}{\sqrt{2\pi}} \exp\left[\mp\left(\frac{\mathrm{i}\pi}{4} + \frac{\pi\varepsilon}{2}\right)\right], \tag{9.15}$$

where $\Gamma(z)$ denotes the Euler gamma function.

Hint: *Make use of the* η*-representation of* $|\Psi_\varepsilon^+\rangle$ *and the relation* $\langle\eta|\xi\rangle = \frac{1}{\sqrt{2\pi}}\exp(-\mathrm{i}\xi\eta)$ *between the states* $|\xi\rangle$ *and* $|\eta\rangle$*, which is a consequence of the commutation relation of the respective operators* $\hat{\xi}$ *and* $\hat{\eta}$*. Moreover, use the definition of the Euler gamma function* $\Gamma(z) = \mathrm{e}^{\mathrm{i}\pi z/2}\int_0^\infty \mathrm{d}x\, x^{z-1}\mathrm{e}^{-\mathrm{i}x}$ *with* $\Gamma(\bar{z}) = \overline{\Gamma(z)}$*, where a bar denotes the complex conjugate quantity.*

g) [*1.5 points*] Finally, we take a closer look at the probability density $|\langle\eta|\Psi_\varepsilon^+\rangle|^2$ for the state $|\Psi_\varepsilon^+\rangle$:

1. Determine the probability density $|\langle\eta|\Psi_\varepsilon^+\rangle|^2$ with the help of the corresponding expressions for the weights $|S_\pm(\varepsilon)|^2$.

   Hint: *Use the relation* $\Gamma\left(\frac{1}{2} - \mathrm{i}\varepsilon\right)\Gamma\left(\frac{1}{2} + \mathrm{i}\varepsilon\right) = \pi/\cosh(\pi\varepsilon)$ *for the Euler gamma function, where* $\cosh(x) = \frac{1}{2}\left(\mathrm{e}^x + \mathrm{e}^{-x}\right)$.

2. Indeed, these weights correspond to the transmission $T(\varepsilon) = |S_-(\varepsilon)|^2$ and reflection coefficient $R(\varepsilon) = |S_+(\varepsilon)|^2$ of the inverted harmonic oscillator. Explain the dependency of these coefficients on the dimensionless energy ε in comparison to the classical situation with the help of Fig. 9.1 and your phase space sketch of part a) and b).

3. Which particular quantum statistics is resembled by the transmission $T(\varepsilon)$ and reflection coefficient $R(\varepsilon)$ of the inverted harmonic oscillator?

h) [*1 point*] In the presence of a black hole, the modes of a scalar quantum field $\Psi_0(t, r)$ with massand spinless quanta can be determined by the Klein-Gordon equation in curved spacetime. For spherical symmetric waves with vanishing angular momentum, the Klein-Gordon equation takes the form

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial r^{*2}}\right) r\Psi_0 = 0 \tag{9.16}$$

in the proximity of the black hole. Here t denotes the time and the radius $r > r_s$, where the Schwarzschild radius $r_s = 2GM/c^2$ is governed by the gravitational constant G, the speed of light c, and the mass M of the black hole. Moreover, we have introduced the Regge-Wheeler tortoise coordinate

$$r^* = r + r_s \ln\left|1 - \frac{r}{r_s}\right|. \tag{9.17}$$

Make use of the separation of variables $\Psi_0(t, r) = q(t)R_0(r)$ in Eq. (9.16) and determine the ordinary differential equations for the functions $q(t)$ in the temporal domain and $rR_0(r)$ in the radial domain. Both equations are connected via the separation constant Ω^2, where the frequency $\Omega = ck$ depends linearly on the respective wave number k. Determine the two linear independent solutions $R_{0,k}^\pm(r)$ of the radial equation for a fixed value of k and show that these modes contain a *logarithmic singularity* that emerges at the event horizon of the black hole located at $r = r_s$.

i) [*1 point*] Central to the emission of Hawking radiation at the event horizon of a black hole is the expansion of a plane wave of frequency ω in the *Kruskal-Szekeres coordinates* u and v, in terms of modes of frequency Ω that display a logarithmic singularity at the event horizon $r = r_s$ and are associated with the *Schwarzschild coordinates* t and r^*. In particular, there exists the decomposition

$$\frac{1}{\sqrt{\omega}}\mathrm{e}^{-\mathrm{i}\omega u} = \int_0^\infty \frac{\mathrm{d}\Omega'}{\sqrt{\Omega'}}\left(\alpha_{\Omega'\omega}\mathrm{e}^{-\mathrm{i}\Omega'\tilde{u}} + \bar{\beta}_{\Omega'\omega}\mathrm{e}^{\mathrm{i}\Omega'\tilde{u}}\right) \tag{9.18}$$

with expansion coefficients $\alpha_{\Omega'\omega}$ and $\beta_{\Omega'\omega}$, where $\tilde{u} = t - r^*$. Moreover, $\bar{\beta}_{\Omega'\omega}$ denotes the complex conjugate of $\beta_{\Omega'\omega}$. Here and in the following we make use of the convention $c = 1$.

1. Show that the expansion coefficient $\beta_{\Omega\omega}$ in Eq. (9.18) can be expressed as

$$\beta_{\Omega\omega} = \frac{1}{2\pi}\sqrt{\frac{\Omega}{\omega}} \int_{-\infty}^{\infty} \mathrm{d}\tilde{u}\, \mathrm{e}^{\mathrm{i}\omega u + \mathrm{i}\Omega\tilde{u}}. \tag{9.19}$$

   Evaluate the integral in Eq. (9.19) for $\Omega > 0$ by using the relation $\tilde{u} = -2r_s \ln\left(-\frac{u}{2r_s}\right)$.

   Hint: *Apply the definition* $\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{\mathrm{i}xy}$ *of the Dirac delta function and recall the definition of the Euler gamma function* $\Gamma(z)$ *provided in part f).*

2. With the help of Eq. (9.19), the mean number of particles emitted by the black hole as Hawking radition in the mode with frequency Ω reads

$$N_\Omega = \int_0^\infty \mathrm{d}\omega\, \left|\beta_{\Omega\omega}\right|^2 = n(\Omega) \int_0^\infty \frac{\mathrm{d}\omega}{2\pi\kappa\omega}. \tag{9.20}$$

   Determine the function $n(\Omega)$ in Eq. (9.20) by using the surface gravity $\kappa = 1/(2r_s)$ of the black hole as a parameter. Compare this result to the one obtained in part g) for the inverted harmonic oscillator.

   Hint: *Apply the identity* $\Gamma(-\mathrm{i}z)\Gamma(\mathrm{i}z) = \pi/\left[z\sinh(\pi z)\right]$ *for the Euler gamma function, where* $\sinh(\pi z) = \frac{1}{2}\left(\mathrm{e}^{\pi z} - \mathrm{e}^{-\pi z}\right)$.

In the end, we want to emphasize that a rigorous treatment of the emission of Hawking radiation at the event horizon of a black hole is only possible within second quantization. However, this exercise gives you some insight into the crucial role of the logarithmic phase singularity for the occurrence of this fascinating phenomenon. Moreover, it reveals the astonishing similarities to the simple quantum system of an inverted harmonic oscillator.

Solution

Further details on the relation between the inverted harmonic oscillator and Hawking radiation emitted by a black hole can be found in the following recently published article:
F. Ullinger, M. Zimmermann, and W.P. Schleich, "The logarithmic phase singularity in the inverted harmonic oscillator", AVS Quantum Sci. **4**, 024402 (2022).

a) [*1.5 points*]

1. In order to identify the respective energy domains in phase space, we make use of the Hamiltonian $H(x,p)$, Eq. (9.2), and the relation $E = H(x_0, p_0)$. Thus, we obtain

   (i) $E < 0: \quad 0 < \frac{p_0^2}{2m} - \frac{1}{2}m\omega^2 x_0^2 \quad \Rightarrow \quad p_0 > \pm m\omega x_0$

   (ii) $E = 0: \quad 0 = \frac{p_0^2}{2m} - \frac{1}{2}m\omega^2 x_0^2 \quad \Rightarrow \quad p_0 = \pm m\omega x_0$

   (iii) $E > 0: \quad 0 = \frac{p_0^2}{2m} - \frac{1}{2}m\omega^2 x_0^2 \quad \Rightarrow \quad p_0 < \pm m\omega x_0$

   The different domains in phase space are sketched in Fig. 9.2 (for simplicity, we have chosen the parameters $m = 1$ and $\omega = 1$). Since the Hamiltonian $H(x,p)$ is time-independent, the energy E is conserved. Thus, classical phase space trajectories $\{x(t), p(t)\}$ with initial position x_0 and momentum p_0 will always stay in the same energy domain of phase space. Consequently, the respective phase space trajectories cannot cross the *horizons* in phase space (thick black lines), which are associated with the energy $E = 0$.

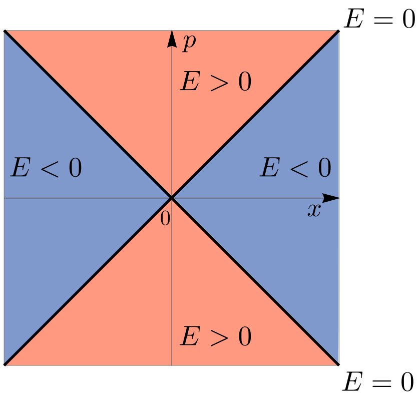


Figure 9.2 Different energy domains for the inverted harmonic oscillator in phase space with position x and momentum p: (i) $E < 0$ (blue), (ii) $E = 0$ (black), (iii) $E > 0$ (red).

2. In order to sketch the trajectories in phase space we recall the Hamiltonian $H(x,p)$, Eq. (9.2), providing the energy

$$E = \frac{p^2}{2m} - \frac{1}{2}m\omega^2 x^2 \tag{9.21}$$

   which is a conserved quantity at any time t. As a consequence, we are able to express the momentum

$$p_\pm(x) = \pm\sqrt{m^2\omega^2 x^2 + 2mE}. \tag{9.22}$$

   as a function of the coordinate x. Consequently, for each energy E, there are two distinct trajectories as shown in Fig. 9.2. Indeed, there we plot the trajectories for the same energies as shown in Fig. 9.1 where this question has already been addressed.

However, the case of a vanishing energy $E = 0$ is nontrivial. Based on Eq. (9.22) one might guess that there are also two trajectories $p_\pm(x) = \pm m\omega x$ corresponding to the two horizons in phase space. However, since a particle traveling along one horizon cannot cross the other horizon, there are four trajectories in this very particular case. Arguably, there is even a fifth one, namely the one corresponding to a particle with $E = 0$ which always remains at the origin of phase space. This issue will become clearer in part b).

Noticing and discussing this subtle point for the energy $E = 0$ will lead to a bonus of 0.5 points. However, with the bonus it is not possible to exceed the maximally achievable points for the entire exercise.

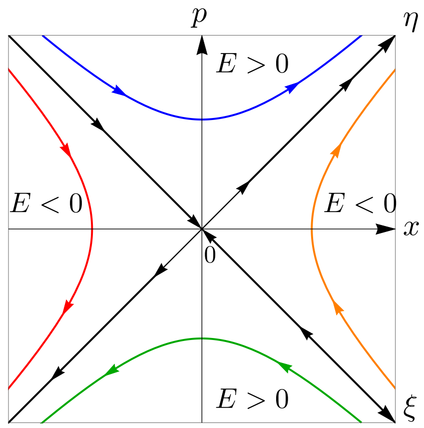


Figure 9.3 The phase space trajectores $\{x(t), p(t)\}$ with the arrow indicating the direction of forward propagation in time t. Two distinct trajectories, corresponding to the same energy E are separated by the two horizons $\xi = 0$ and $\eta = 0$, depicted along the diagonals, which divide phase space into four disjunct regions. On the half plane $\eta < 0$, the phase space trajectories (blue and red) describe the motion of a particle that approaches the potential barrier from the left, while for $\eta > 0$ the phase space trajectories (green and orange) correspond to an incoming particle from the right. In the domain $\eta\xi < 0$ the trajectories (blue and green) belong to a particle with positive energy $E > 0$, while in the domain $\eta\xi > 0$, the trajectories (red and orange) correspond to a particle with negative energy $E < 0$. A particle with energy $E = 0$ travels along the horizons $\eta = 0$ or $\xi = 0$, as indicated by the black arrows. Figure reprinted from F. Ullinger et al., AVS Quantum Sci. **4**, 024402 (2022) under license CC BY 4.0.

b) [*1.5 points*]

1. With the help of the Hamiltonian H, Eq. (9.5), the equations of motion given by Eqs. (9.6) and (9.7) take the form

$$\dot{\xi} = -\omega\xi, \tag{9.23}$$
$$\dot{\eta} = \omega\eta \tag{9.24}$$

with the solutions

$$\xi(t) = \xi_0 \mathrm{e}^{-\omega t}, \tag{9.25}$$
$$\eta(t) = \eta_0 \mathrm{e}^{\omega t}. \tag{9.26}$$

Thus, we obtain with $\xi_1 = \xi(T_1)$ the time

$$T_1 = \frac{1}{\omega} \ln\left(\frac{\xi_0}{\xi_1}\right), \tag{9.27}$$

which is governed by a logarithmic dependence.

2. From Eq. (9.27) we observe that in the limit $\xi_1 \to 0$ the time $T_1 \to \infty$ in order to reach the horizon $\xi = 0$. Indeed, for $\xi_0 < 0$ we consider the limit $\xi_1 \to 0-$ and for $\xi_0 > 0$, we consider the limit $\xi_1 \to 0+$ and recall that $\omega > 0$. Thus, the horizon $\xi = 0$ cannot be reached within a finite time from *any* initial coordinate $\xi_0 \neq 0$, that is from any point in phase space which does not lie on the horizon $\xi = 0$. Obviously, this result is independent of the value of η_0 (and also valid for $\eta_0 = 0$, that is the particle travels along the other horizon, see also the black lines in Fig. 9.3).

   As evident from the Hamiltonian H, Eq. (9.5), a particle traveling along the horizons $\xi = 0$ or $\eta = 0$ possesses the energy $E = 0$. Thus, this result also gives a more detailed explanation for the issue with the number of possible trajectories of energy $E = 0$ which was raised in part a). Indeed, no classical particle can cross one of the two horizons in phase space in a finite amount of time.

3. The horizons in phase space are determined by the conditions $\xi = 0$ and $\eta = 0$, or equivalently, $p = \pm m\omega x$ according to Eqs. (9.3) and (9.4). In the limit $\omega \to 0$, we thus obtain $p = 0$. Accordingly, both horizons align with the axis of vanishing momentum. As a consequence, the energy domains with $E < 0$ disappear and only energies $E \geq 0$ are allowed. Indeed, this result is reasonable since for $\omega \to 0$ we obtain a free particle with energy $E \geq 0$. The horizons and the energy domains in phase in the limit $\omega \to 0$ are sketched in Fig. 9.4.

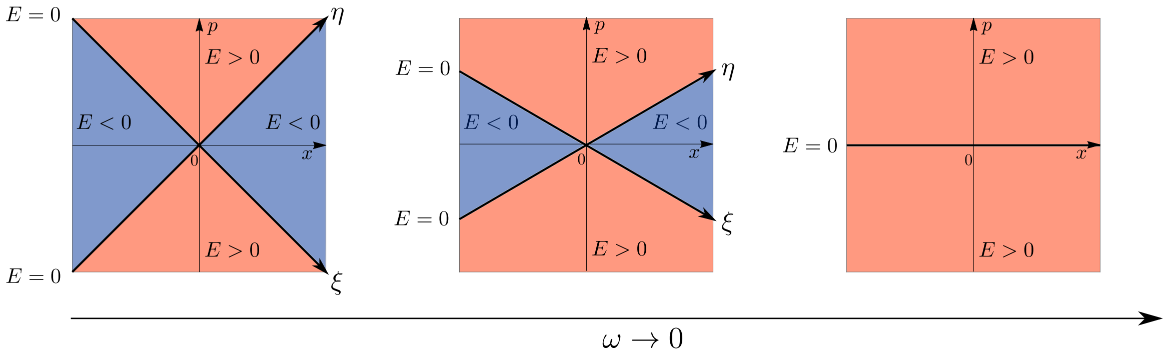


Figure 9.4 Change of the energy domains of the inverted harmonic oscillator in phase space as $\omega \to 0$.

c) [*0.5 points*] The commutation relation for the operators $\hat{\xi}$ and $\hat{\eta}$ reads

$$\left[\hat{\xi}, \hat{\eta}\right] = \frac{m\omega}{2\hbar}\left[\hat{x} - \frac{\hat{p}}{m\omega}, \hat{x} + \frac{\hat{p}}{m\omega}\right] \tag{9.28}$$

$$= \frac{m\omega}{2\hbar}\left(\frac{1}{m\omega}[\hat{x}, \hat{p}] - \frac{1}{m\omega}[\hat{p}, \hat{x}]\right) \tag{9.29}$$

$$= \frac{1}{\hbar}[\hat{x}, \hat{p}] \tag{9.30}$$

$$= \mathrm{i}, \tag{9.31}$$

where we have used the standard commutation relation $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$. Consequently, $\hat{\xi}$ and $\hat{\eta}$ are *canonically conjugate operators*, analogous to the position operator $\hat{x}$ and the momentum operator $\hat{p}$. Thus, there exists an uncertainty relation with regard to these operators.

d) [*1 point*] We consider the stationary Schrödinger equation

$$-\frac{\hbar\omega}{2}\left(\hat{\xi}\hat{\eta} + \hat{\eta}\hat{\xi}\right)|\varepsilon\rangle = \hbar\omega\,|\varepsilon\rangle\,. \tag{9.32}$$

Since $\hat{\xi}$ and $\hat{\eta}$ are canonically conjugate operators with $[\hat{\xi},\hat{\eta}]=\mathrm{i}$, see part c), we can use the replacements $\hat{\xi}\to\xi$ and $\hat{\eta}\to-\mathrm{i}\frac{\partial}{\partial\xi}$ when working in ξ-representation. The latter expression is analogous to the position-representation of the momentum operator. As a consequence, Eq. (9.32) reads

$$\frac{\hbar\omega\mathrm{i}}{2}\left(\xi\frac{\mathrm{d}}{\mathrm{d}\xi}+\frac{\mathrm{d}}{\mathrm{d}\xi}\xi\right)\langle\xi|\varepsilon\rangle=\hbar\omega\varepsilon\langle\xi|\varepsilon\rangle\,. \tag{9.33}$$

Next, we recast Eq. (9.33) as

$$\xi\frac{\mathrm{d}}{\mathrm{d}\xi}\langle\xi|\varepsilon\rangle=\left(-\frac{1}{2}-\mathrm{i}\varepsilon\right)\langle\xi|\varepsilon\rangle\,. \tag{9.34}$$

First, we notice from the left-hand side of Eq. (9.34) that there is a problem when $\xi\to0$ (that is when the horizon is approached). Thus, Eq. (9.34) can be solved separately in the domains $\xi<0$ and $\xi>0$. For this purpose, we recast $\Psi_\varepsilon^\pm(\xi)$, given by Eq. (9.12), as

$$\Psi_\varepsilon^\pm(\xi)=\frac{1}{\sqrt{2\pi}}\,|\xi|^{-1/2-\mathrm{i}\varepsilon}\,\Theta(\pm\xi) \tag{9.35}$$

Thus, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\Psi_\varepsilon^\pm(\xi)=\begin{cases}\frac{1}{\sqrt{2\pi}}\left[(-\frac{1}{2}-\mathrm{i}\varepsilon)\xi^{-3/2-\mathrm{i}\varepsilon}\Theta(\pm\xi)\right], & \xi>0,\\ \frac{1}{\sqrt{2\pi}}\left[-(-\frac{1}{2}-\mathrm{i}\varepsilon)(-\xi)^{-3/2-\mathrm{i}\varepsilon}\Theta(\pm\xi)\right], & \xi<0.\end{cases} \tag{9.36}$$

Consequently, we arrive at

$$\xi\frac{\mathrm{d}}{\mathrm{d}\xi}\Psi_\varepsilon^\pm(\xi)=\left(-\frac{1}{2}-\mathrm{i}\varepsilon\right)\frac{1}{\sqrt{2\pi}}\,|\xi|^{-1/2-\mathrm{i}\varepsilon}\,\Theta(\pm\xi) \tag{9.37}$$

for $\xi\neq0$ and thus, Eq. (9.35), constitutes a solution of Eq. (9.34). However, due to the singularity, this wave function is not continuous at the horizon $\xi=0$.

e) [*0.5 points*] In order to determine the energy eigenstates $|\Phi_\varepsilon^\pm\rangle$ we recall the commutator $\left[\hat{\xi},\hat{\eta}\right]=\mathrm{i}$. Thus, the stationary Schrödinger equation (9.32) reads in η-representation

$$-\frac{\hbar\omega\mathrm{i}}{2}\left(\eta\frac{\mathrm{d}}{\mathrm{d}\eta}+\frac{\mathrm{d}}{\mathrm{d}\eta}\eta\right)\langle\eta|\varepsilon\rangle=\hbar\omega\varepsilon\langle\eta|\varepsilon\rangle\,. \tag{9.38}$$

Except of the minus sign in front of the right-hand side, Eq. (9.38) has the same form as Eq. (9.33). By recalling the solutions $\Psi_\varepsilon^\pm(\xi)$, Eq. (9.12), of Eq. (9.34) we identify the solutions

$$\Phi_\varepsilon^\pm(\eta)=\frac{1}{\sqrt{2\pi\,|\eta|}}\exp\left(\mathrm{i}\varepsilon\ln|\eta|\right)\Theta\left(\pm\eta\right) \tag{9.39}$$

of Eq. (9.38), which are proportional to $\Theta(\pm\eta)$ and where the opposite sign appears in the front of the logarithmic phase.

As consequence, we observe that there exist two sets of energy eigenstates $|\Psi_\varepsilon^\pm\rangle$ and $|\Phi_\varepsilon^\pm\rangle$ with energy ε, whose ξ- and η-representation have the same form except of a sign difference. This feature is a consequence of the invariance of the Hamiltonian $\hat{H}$, Eq. (9.10), with regard to an exchange of the operators $\hat{\xi}$ and $\hat{\eta}$.

f) [*1.5 points*] In order to express the state $|\Psi_\varepsilon^+\rangle$ as a superposition of the states $|\Phi_\varepsilon^+\rangle$ and $|\Phi_\varepsilon^-\rangle$, we consider its η-representation and insert the identity in terms of the states $|\xi\rangle$:

$$\langle\eta|\Psi_\varepsilon^+\rangle = \int_{-\infty}^{\infty} \mathrm{d}\xi\, \langle\eta|\xi\rangle \langle\xi|\Psi_\varepsilon^+\rangle \tag{9.40}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}\xi\, \frac{1}{\sqrt{2\pi}} \exp(-\mathrm{i}\xi\eta) \frac{1}{\sqrt{2\pi\,|\xi|}} \exp\left(-\mathrm{i}\varepsilon \ln|\xi|\right) \Theta(\xi) \tag{9.41}$$

$$= \frac{1}{2\pi} \int_0^{\infty} \mathrm{d}\xi\, \xi^{-1/2-\mathrm{i}\varepsilon} \mathrm{e}^{-\mathrm{i}\xi\eta}, \tag{9.42}$$

where we have made use of Eqs. (9.12).

Next, we apply the substitution $x = \xi\eta$ and arrive at

$$\langle\eta|\Psi_\varepsilon^+\rangle = \frac{1}{2\pi}\left[\Theta(\eta)\int_0^{\infty} \mathrm{d}x\, \eta^{-1}\left(\frac{x}{\eta}\right)^{-1/2-\mathrm{i}\varepsilon} \mathrm{e}^{-\mathrm{i}x} + \Theta(-\eta)\int_0^{-\infty} \mathrm{d}x\, \eta^{-1}\left(\frac{x}{\eta}\right)^{-1/2-\mathrm{i}\varepsilon} \mathrm{e}^{-\mathrm{i}x}\right] \tag{9.43}$$

$$= \frac{1}{2\pi}\left[\eta^{-1/2+\mathrm{i}\varepsilon}\Theta(\eta)\int_0^{\infty} \mathrm{d}x\, x^{-1/2-\mathrm{i}\varepsilon}\mathrm{e}^{-\mathrm{i}x} + (-\eta)^{-1/2+\mathrm{i}\varepsilon}\Theta(-\eta)\int_0^{\infty} \mathrm{d}x\, x^{-1/2-\mathrm{i}\varepsilon}\mathrm{e}^{\mathrm{i}x}\right] \tag{9.44}$$

$$= \frac{1}{2\pi}\Gamma\left(\frac{1}{2}-\mathrm{i}\varepsilon\right)|\eta|^{-1/2+\mathrm{i}\varepsilon}\left[\mathrm{e}^{-\mathrm{i}\pi(1/2-\mathrm{i}\varepsilon)/2}\Theta(\eta) + \mathrm{e}^{+\mathrm{i}\pi(1/2-\mathrm{i}\varepsilon)}\Theta(-\eta)\right] \tag{9.45}$$

$$= \frac{\Gamma\left(\frac{1}{2}-\mathrm{i}\varepsilon\right)}{\sqrt{2\pi}}\mathrm{e}^{-\mathrm{i}\pi/4-\pi\varepsilon/2}\langle\eta|\Phi_\varepsilon^+\rangle + \frac{\Gamma\left(\frac{1}{2}-\mathrm{i}\varepsilon\right)}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}\pi/4+\pi\varepsilon/2}\langle\eta|\Phi_\varepsilon^-\rangle\,, \tag{9.46}$$

where we have used the definition of the Euler gamma function $\Gamma(z)$.

g) [*1.5 points*]

1. With the help of Eq. (9.14), we obtain

$$\left|\langle\eta|\Psi_\varepsilon^+\rangle\right|^2 = \langle\eta|\Psi_\varepsilon^+\rangle\langle\Psi_\varepsilon^+|\eta\rangle \tag{9.47}$$

$$= |S_+(\varepsilon)|^2\left|\langle\eta|\Phi_\varepsilon^+\rangle\right|^2 + |S_-(\varepsilon)|^2\left|\langle\eta|\Phi_\varepsilon^-\rangle\right|^2, \tag{9.48}$$

   where we have made use of the fact that $\langle\eta|\Phi_\varepsilon^+\rangle\langle\Phi_\varepsilon^-|\eta\rangle = \langle\eta|\Phi_\varepsilon^-\rangle\langle\Phi_\varepsilon^+|\eta\rangle = 0$. By using Eq. (9.15), we can show that

$$|S_\pm(\varepsilon)|^2 = \frac{1}{2\pi}\Gamma\left(\frac{1}{2}-\mathrm{i}\varepsilon\right)\Gamma\left(\frac{1}{2}+\mathrm{i}\varepsilon\right)\exp\left(\mp\pi\varepsilon\right) \tag{9.49}$$

$$= \frac{1}{2\pi}\frac{\pi}{\cosh(\pi\varepsilon)}\exp\left(\mp\pi\varepsilon\right) \tag{9.50}$$

$$= \frac{1}{1+\mathrm{e}^{\pm 2\pi\varepsilon}}. \tag{9.51}$$

   As a result, we arrive at

$$\left|\langle\eta|\Psi_\varepsilon^+\rangle\right|^2 = \frac{1}{2\pi\,|\eta|}\left[\frac{\Theta(-\eta)}{1+\mathrm{e}^{-2\pi\varepsilon}} + \frac{\Theta(\eta)}{1+\mathrm{e}^{2\pi\varepsilon}}\right]. \tag{9.52}$$

2. A classical particle incoming from the left is reflected at the potential barrier of the inverted harmonic oscillator for an energy $E < 0$, while it is transmitted for an energy $E > 0$. This feature is evident from both Fig. 9.1 and the phase space trajectories. In contrary, a quantum

particle can tunnel through the potential well, while its transmission $T(\varepsilon)$ and reflection coefficients $R(\varepsilon)$ are smooth functions of the energy ε. For $\varepsilon = 0$, we obtain $T(0) = R(0) = 1/2$, that is transmission and reflection are equally likely. Only for very large energies $\varepsilon \to \pm\infty$ we recover the classical situation: $T(\varepsilon \to +\infty) = 1$ and $R(\varepsilon \to +\infty) = 0$ as well as $T(\varepsilon \to -\infty) = 0$ and $R(\varepsilon \to -\infty) = 1$.

3. Both the transmission and reflection coefficients of the inverted harmonic oscillator resemble the *Fermi-Dirac statistics*
$$f(E) = \frac{1}{1 + \mathrm{e}^{(E-\mu)/k_\mathrm{B}T}} \tag{9.53}$$
with the total chemical potential μ. In contrast to thermodynamics where the characteristic energy of the system is given by the product $k_\mathrm{B}T$ of the Boltzmann constant k_B and the temperature T, it is the product $h\omega$ of the Planck constant h and the steepness ω divided by $4\pi^2$, which defines the energy scale of the inverted harmonic oscillator in the expression
$$|S_\pm(\varepsilon)|^2 = \frac{1}{1 + \mathrm{e}^{\pm 2\pi\varepsilon}} \tag{9.54}$$
(remember that the energy was defined as $\hbar\omega\varepsilon$). This particular feature is a consequence of the Fourier transform of a logarithmic phase singularity in combination with an amplitude singularity.

h) [*1 point*] We apply the separation of variables $\Psi_0(t, r) = q(t)R_0(r)$ and recast Eq. (9.16) in the form
$$\frac{\partial_t^2 q(t)}{q(t)} = c^2 \frac{\partial_{r^*}^2 \left[rR_0(r)\right]}{rR_0(r)}. \tag{9.55}$$
Since the left-hand side only depends on time t and the right hand-side only depends on the radius r, and t and r are indpendent variables, both sides are equal to a separation constant which we call Ω^2. As a consequence, we obtain in the temporal domain the oscillator equation
$$\left(\frac{\partial^2}{\partial t^2} + \Omega^2\right) q(t) = 0 \tag{9.56}$$
and in the radial direction, we arrive at the differential equation
$$\left(\frac{\partial^2}{\partial r^{*2}} + k^2\right) rR_0(r) = 0 \tag{9.57}$$
with $k = \Omega/c$.

We focus on the radial equation (9.57) for which we identify the solutions
$$R_{0,k}^\pm(r) = \frac{\mathrm{e}^{\pm \mathrm{i}kr^*}}{r} = \frac{1}{r}\mathrm{e}^{\pm \mathrm{i}k\left(r + r_s \ln\left|1 - \frac{r}{r_s}\right|\right)}, \tag{9.58}$$
where we have made use of the definition of r^* given by Eq. (9.17). The mode functions $R_{0,k}^\pm(r)$ display a *logarithmic phase singularity* at the event horizon of the black hole located at $r = r_s$. In contrast to the energy eigenstates of the inverted harmonic oscillator, the functions $R_{0,k}^\pm(r)$ only display a *phase singularity* and no *amplitude singularity* at $r = r_s$. This difference has dramatic consequences as shown in part i).

In addition, the occupation of a particular mode can be analyzed by quantizing the temporal Eq. (9.56), leading to the appearance of creation and annihilation operators which are familar from the harmonic oscillator.

i) [*1 point*]

1. First, we multiply the given decomposition by the factor $\exp(-\mathrm{i}\Omega\tilde{u})$ and obtain

$$\frac{1}{\sqrt{\omega}}\mathrm{e}^{-\mathrm{i}\omega u}\mathrm{e}^{-\mathrm{i}\Omega\tilde{u}} = \int_0^\infty \frac{\mathrm{d}\Omega'}{\sqrt{\Omega'}}\left[\alpha_{\Omega'\omega}\mathrm{e}^{-\mathrm{i}(\Omega'+\Omega)\tilde{u}} + \bar{\beta}_{\Omega'\omega}\mathrm{e}^{\mathrm{i}(\Omega'-\Omega)\tilde{u}}\right]. \tag{9.59}$$

Next, integrating both sides of the equation with regard to $\tilde{u}$ yields

$$\frac{1}{\sqrt{\omega}}\int_{-\infty}^\infty \mathrm{d}\tilde{u}\,\mathrm{e}^{-\mathrm{i}\omega u}\mathrm{e}^{-\mathrm{i}\Omega\tilde{u}} = 2\pi\int_0^\infty \frac{\mathrm{d}\Omega'}{\sqrt{\Omega'}}\left[\alpha_{\Omega'\omega}\delta(\Omega'+\Omega) + \bar{\beta}_{\Omega'\omega}\delta(\Omega'-\Omega)\right]. \tag{9.60}$$

By evaluating the right-hand side for $\Omega > 0$ we arrive at

$$\frac{1}{\sqrt{\omega}}\int_{-\infty}^\infty \mathrm{d}\tilde{u}\,\mathrm{e}^{-\mathrm{i}\omega u}\mathrm{e}^{-\mathrm{i}\Omega\tilde{u}} = \frac{2\pi}{\sqrt{\Omega}}\bar{\beta}_{\Omega\omega} \tag{9.61}$$

or, equivalently,

$$\beta_{\Omega\omega} = \frac{1}{2\pi}\sqrt{\frac{\Omega}{\omega}}\int_{-\infty}^\infty \mathrm{d}\tilde{u}\,\mathrm{e}^{\mathrm{i}\omega u+\mathrm{i}\Omega\tilde{u}}. \tag{9.62}$$

Next, the relation $\tilde{u} = -2r_s\ln\left(-\frac{u}{2r_s}\right)$ yields

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}u} = -\frac{2r_s}{u} \tag{9.63}$$

and Eq. (9.62) takes the form

$$\beta_{\Omega\omega} = \frac{1}{2\pi}\sqrt{\frac{\Omega}{\omega}}\int_{-\infty}^0 \mathrm{d}u\,\mathrm{e}^{\mathrm{i}\omega u}\mathrm{e}^{-2\mathrm{i}r_s\Omega\ln\left(-\frac{u}{2r_s}\right)}\left(-\frac{2r_s}{u}\right) \tag{9.64}$$

$$= \frac{1}{2\pi\omega}\sqrt{\frac{\Omega}{\omega}}\int_0^\infty \mathrm{d}s\,\mathrm{e}^{-\mathrm{i}s}\left(\frac{s}{2r_s\omega}\right)^{-1-2\mathrm{i}r_s\Omega} \tag{9.65}$$

with $s = -\omega u$. By recalling the definition of the Euler gamma function $\Gamma(z)$ we finally recast Eq. (9.65) as

$$\beta_{\Omega\omega} = \frac{r_s}{\pi}\sqrt{\frac{\Omega}{\omega}}\left(2r_s\omega\right)^{2\mathrm{i}r_s\Omega}\mathrm{e}^{-\pi r_s\Omega}\Gamma\left(-2\mathrm{i}r_s\Omega\right). \tag{9.66}$$

Indeed, we observe that it is again the Fourier transform of a logarithmic phase singularity which leads to the appearance of the Euler gamma function.

2. First, we make use of our previous result, Eq. (9.66), and obtain

$$\int_0^\infty \mathrm{d}\omega\ |\beta_{\Omega\omega}|^2 = \frac{r_s^2}{\pi^2}\int_0^\infty \mathrm{d}\omega\left(\frac{\Omega}{\omega}\right)\mathrm{e}^{-2\pi r_s\Omega}\Gamma\left(-2\mathrm{i}r_s\Omega\right)\Gamma\left(2\mathrm{i}r_s\Omega\right). \tag{9.67}$$

Next, we apply the identity $\Gamma(-\mathrm{i}z)\Gamma(\mathrm{i}z) = \pi/\left[z\sinh(\pi z)\right]$ and find

$$\int_0^\infty \mathrm{d}\omega\ |\beta_{\Omega\omega}|^2 = \frac{r_s^2\Omega}{\pi^2}\frac{\pi}{\Omega r_s}\frac{\mathrm{e}^{-2\pi r_s\Omega}}{\mathrm{e}^{2\pi r_s\Omega}-\mathrm{e}^{-2\pi r_s\Omega}}\int_0^\infty \mathrm{d}\omega\,\omega^{-1} \tag{9.68}$$

$$= \frac{1}{\mathrm{e}^{4\pi r_s\Omega}-1}\int_0^\infty \mathrm{d}\omega\,\frac{r_s}{\pi\omega}. \tag{9.69}$$

By using the definition for the surface gravity $\kappa = 1/(2r_s)$ of the black hole, we arrive at the familiar expression

$$n(\Omega) = \frac{1}{\mathrm{e}^{2\pi\Omega/\kappa} - 1} \tag{9.70}$$

for the mean density of particles with frequency Ω emitted by a black hole via Hawking radiation. As shown in Eq. (9.70), the function $n(\Omega)$ is governed by a *Bose-Einstein distribution*.

In the case of the inverted harmonic oscillator, we have obtained a Fermi-Dirac distribution as a consequence of the Fourier transform of a logarithmic phase singularity in combination with an inverse square-root amplitude singularity. This particular dependency is reflected in the transmission and reflection coefficients of this system. Contrarily, in our discussion of the black hole it was the Fourier transform of a logarithmic phase singularity with a simple pole in the amplitude which lead to the appearance of a Bose-Einstein distribution.

Problem 10

Rare and Extreme Events in Nonlinear Physics: From Fiber Optics to Oceanic Waves

Prof. Dr. Christophe Finot – *University of Burgundy*

Background Single mode optical fibers are a nearly ideal medium of propagation: the very low level of attenuation of silica combined with the single mode behavior of the fiber enables light to propagate with a spatial profile unaffected over kilometers and kilometers. However, in the case of ultrashort and powerful pulses, two effects have to be taken into account: the dispersion and the nonlinearity that are related to the dependence of the optical index on the frequency and power, respectively.

In the present exercise, we explore the intriguing relationship between the propagation of light in single mode optical fibers and the occurrence of rare and extreme events for oceanic waves.

In the first part of the exercise, we compare the propagation of ultrashort optical pulses and the spatial profile of continuous electromagnetic waves. For an ultrashort optical pulse we display in Fig. 10.1 (a) the intensity of the electromagnetic field including the optical carrier (blue) and the temporal profile of the slowly-varying envelope $|\psi(z,t)|^2$ (red) at a given position $z = z_0$. In Fig. 10.1 (b) we show the spatial intensity distribution $|\Psi(z,x)|^2$ of a continuous electromagnetic wave as a function of the propagation coordinate z and the transversal coordinate x.

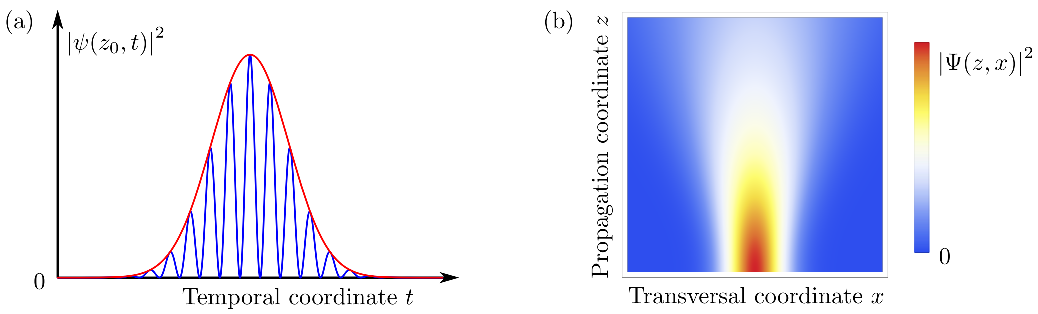


Figure 10.1 (a) The temporal intensity profile of an ultrashort optical pulse (blue) and its envelope $|\psi(z,t)|^2$ (red) at a given position $z = z_0$. (b) The spatial intensity profile $|\Psi(z,x)|^2$ of a continuous electromagnetic wave with regard to the transversal coordinate x and the propagation coordinate z.

a) [*0.5 points*] We study the dynamics of an ultrashort pulse with duration above 1 ps as depicted in Fig. 10.1 (a). Under the approximation of a slowly-varing envelope (i.e. the optical carrier of the light pulse is neglected) and within the scalar approximation, the evolution of the complex electric field $\psi(z,t)$ is governed by the partial differential equation

$$\mathrm{i}\frac{\partial \psi}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \psi}{\partial t^2} \tag{10.1}$$

with z and t being the distance and temporal coordinate, respectively. Here β_2 denotes the second-order dispersion coefficient. At the position $z = z_0$, the shape of the pulse is determined by the initial condition $\psi(z_0,t) = \psi_0(t)$.

In order to understand the effect of dispersion, it is beneficial to work in Fourier space. The Fourier transform of $\psi(z,t)$ is given by

$$\tilde{\psi}(z,\omega) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \psi(z,t). \tag{10.2}$$

Using Eq. (10.1) please derive a solution for $\tilde{\psi}(z,\omega)$ and explain the effect of dispersion with it.

b) [*0.5 points*] We turn now to the spatial propagation of the complex electric field $\Psi(x,z)$ of a continuous wave as displayed in Fig. 10.1 (b). The Helmholtz equation that rules diffraction under 1D paraxial conditions reads

$$\mathrm{i}\frac{\partial \Psi}{\partial z} = -\frac{\lambda}{4\pi}\frac{\partial^2 \Psi}{\partial x^2} \tag{10.3}$$

with z and x being the longitudinal and transverse coordinate, respectively. Here λ denotes the wavelength and $\Psi_0(x) = \Psi(z_0, x)$ is the initial profile at $z = z_0$.

What are the analogies that can be drawn between the dispersion of an ultrashort pulse and the diffraction of a continuous wave? Comment on the impact of dispersion and diffraction.

Hint: *Make use of Fourier space for your analysis.*

c) [*0.5 points*] Next, we analyze the dynamics of an ultrashort optical pulse $\psi(z,t)$ for a negligible second-order dispersion coefficient $|\beta_2| \ll 1$ in Eq. (10.1). When the effects of nonlinearity in the optical fibre become significant, the electromagnetic field evolves according to the equation

$$\frac{\partial \psi}{\partial z} = \mathrm{i}\gamma \left|\psi\right|^2 \psi, \tag{10.4}$$

where γ is the nonlinear coefficient of the optical fiber.

Consider the case that the intensity profile $|\psi(z,t)|^2$ is independent of the distance z and determine the corresponding solution of Eq. (10.4). What is the physical consequence of the so-called Kerr nonlinearity in Eq. (10.4) on the temporal profile of the pulse?

d) [*1 point*] For an ultrashort optical pulse in a nonlinear fiber, an effect called self-focusing can occur. To obtain a deeper insight into this phenomenon, we establish again an analogy to the spatial profile $\Psi(z,x)$ of a continuous electromagnetic wave. In wave optics, the impact of a perfectly converging 1D lens of focal length f situated at the propagation coordinate $z = z_1$ can be described as

$$\Psi(z_1^+, x) = \exp\left(-\mathrm{i}\frac{\pi x^2}{\lambda f}\right)\Psi(z_1^-, x), \tag{10.5}$$

where $\Psi(z_1^-, x)$ and $\Psi(z_1^+, x)$ are the fields directly before and after the lens. Here λ is the wavelength introduced in Eq. (10.3).

Explain why the nonlinearity in Eq. (10.4) acts like a converging lens. For this purpose, we analyze an ultrashort optical pulse with a symmetric temporal profile $\psi(z_0, t)$ at $z = z_0$. As an example please use a Gaussian pulse $\psi(z,t) = \psi_{\max}\mathrm{e}^{-\frac{t^2}{K^2}}$, with temporal width K and amplitude $\psi_{\max}$. For this exercise it is sufficient to approximate the pulse shape to second-order in t. Note, that the intensity profile $|\psi(z,t)|^2$ remains independent of z as in c). Determine the waveform $\psi(z = d, t)$ at the position $z = d > z_0$ and compare it to Eq. (10.5).

In the second part of the exercise, we take a closer look at nonlinear effects. Indeed, for ultra-short high-power waveforms, both dispersive and nonlinear effects act simultaneously. Thus, the evolution of the optical pulse is described by the nonlinear Schrödinger equation

$$\mathrm{i}\frac{\partial \psi}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \psi}{\partial t^2} - \gamma \left|\psi\right|^2 \psi. \tag{10.6}$$

By using the normalized coordinates $\xi = z/L$ and $\tau = t/t_0$ with the characteristic length $L = 1/(\gamma P_0)$ and time $t_0 = \sqrt{|\beta_2|\, L}$, Eq. (10.6) reduces for $\beta_2 < 0$ to

$$\mathrm{i}\frac{\partial A}{\partial \xi} = -\frac{1}{2}\frac{\partial^2 A}{\partial \tau^2} - \left|A\right|^2 A, \tag{10.7}$$

where the function $A = \psi/\sqrt{P_0}$ contains the typical power P_0 of the waveform under study.

e) [*2 points*] The mathematical solutions of the nonlinear Eq. (10.7) are non-trivial. However, some specific wave forms have been identified in the past. We focus our attention on two of these waves, the soliton

$$A_s(\xi, \tau) = \operatorname{sech}(\tau)\mathrm{e}^{\mathrm{i}\xi/2} \tag{10.8}$$

and the Peregrine breather

$$A_p(\xi, \tau) = \left[1 - \frac{4\,(1 + 2\mathrm{i}\xi)}{1 + 4\tau^2 + 4\xi^2}\right]\mathrm{e}^{\mathrm{i}\xi}. \tag{10.9}$$

Here we have introduced the hyperbolic secant function

$$\operatorname{sech}(\tau) = \frac{2}{\mathrm{e}^{\tau} + \mathrm{e}^{-\tau}}. \tag{10.10}$$

Plot the very approximate intensity profiles $|A_s|^2$ and $|A_p|^2$ at $\xi = 0$, $\xi = -\infty$ and $\xi = \infty$ as a function of τ. Identify in both cases the value of the intensity at $\tau = \pm\infty$ as well as the peak power at $\tau = 0$. How does the peak power of the two waves evolve as a function of ξ? Identify the main differences between these two waves.

Hint: *For the soliton case, you may find insights by using a rough approximation of the exponential to the second-order; qualitative links to a Lorentzian waveform can also be drawn.*

f) [*1 point*] All the effects observed in e) are intimately linked to a process that is called modulation instability. In order to highlight this phenomenon, let us switch again to the unnormalized Schrödinger equation (10.6) and consider a continuous wave

$$\psi_C(z, t) = \psi_0 \exp\left(\mathrm{i}\gamma P_0 z\right) \tag{10.11}$$

which is a solution of equation (10.6). Here ψ_0 and P_0 are real-valued and correspond to the initial wave amplitude and power, respectively. In order to check if the solution $\psi_C(z, t)$, Eq. (10.11), is stable against perturbation, we consider the perturbed wave

$$\psi_\varepsilon(z, t) = \left[\psi_0 + \varepsilon(z, t)\right]\exp\left(\mathrm{i}\gamma P_0 z\right) = \psi_C + \varepsilon(z, t)\exp\left(\mathrm{i}\gamma P_0 z\right). \tag{10.12}$$

By plugging the perturbed wave $\psi_\varepsilon(z,t)$ into the nonlinear Schrödinger equation (10.6), derive the equation

$$\mathrm{i}\frac{\partial \varepsilon}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \varepsilon}{\partial t^2} - \gamma P_0 \left(\varepsilon + \varepsilon^*\right) \tag{10.13}$$

that is satisfied by the function $\varepsilon(z,t)$.

Hint: *The perturbation is assumed to be small, so that terms of order ε^2 can be neglected.*

g) [*2 points*] We are looking for a solution of Eq. (10.13) that can be expressed as a plane wave

$$\varepsilon(z,t) = a_1 \exp\left[\mathrm{i}(kz-\omega t)\right] + a_2 \exp\left[-\mathrm{i}\left(k^* z - \omega t\right)\right], \tag{10.14}$$

where a_1, a_2, and k are complex-valued constants. Show that a_1 and a_2 are solutions of the system

$$\left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 - k\right) a_1 + \gamma P_0 a_2^* = 0$$

$$\gamma P_0 a_1^* + \left(\frac{\beta}{2}\omega^2 + \gamma P_0 + k^*\right) a_2 = 0 \tag{10.15}$$

of coupled equations. For certain k these coupled equations can have non-trivial solutions. Determine k for a non-trivial solution as a function of the parameters β_2, ω, γ, and P_0.

h) [*1.5 points*] The nonlinear system will be stable against perturbation when k is purely real. Otherwise, when k contains an imaginary part, an exponential growth of the perturbation will happen as evident from Eq. (10.14). Can the propagation be unstable in the regime of normal dispersion ($\beta_2 > 0$) or anormalous dispersion ($\beta_2 < 0$)? Derive the range of the frequency ω for which the wave is unstable. What is the frequency for which most of the gain is observed (the gain is given by the imaginary part of k). Draw the approximate shape of the gain as a function of the frequency ω.

i) [*1 point*] The nonlinear Schrödinger equation also models the propagation of oceanic waves. In this context, the envelope $u(z,t)$ of a modulated wave train with wave vector k_0 propagating in a 1D water tank along the z-direction is governed by the equation

$$\mathrm{i}\frac{\partial u}{\partial z} = \frac{1}{g}\frac{\partial^2 u}{\partial t^2} + k_0^3 \left|u\right|^2 u \tag{10.16}$$

with g being the gravitational acceleration. By using normalized quantities, Eq. (10.16) can be cast into the form

$$\mathrm{i}\frac{\partial U}{\partial \xi} = -\frac{1}{2}\frac{\partial^2 U}{\partial \tau^2} - \left|U\right|^2 U. \tag{10.17}$$

Propose a solution $U(\xi,\tau)$ of Eq. (10.17) that could be relevant for the explanation of oceanic rogue waves. These are particular waves that appear from nowhere and disappear without leaving a trace.

Solution

More information on the solution of the nonlinear Schrödinger equation with dispersion for part d)-i) can be found here: `https://www.nature.com/articles/nphys1740.pdf`

a) The inverse Fourier transform is given by

$$\psi(z,t) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t} \tilde{\psi}(z,\omega). \tag{10.18}$$

We insert this expression into both sides of Eq. (10.1) and obtain

$$\mathrm{i}\frac{\partial}{\partial z}\left[\frac{1}{\sqrt{2\pi}} \int \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t} \tilde{\psi}(z,\omega)\right] = \frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}\left[\frac{1}{\sqrt{2\pi}} \int \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t} \tilde{\psi}(z,\omega)\right] \tag{10.19}$$

$$0 = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega t}\left[\frac{\partial}{\partial z}\tilde{\psi}(z,\omega) - \mathrm{i}\omega^2\frac{\beta_2}{2}\tilde{\psi}(z,\omega)\right]. \tag{10.20}$$

This is solved by

$$\tilde{\psi}(z,\omega) = \tilde{\psi}_0(\omega)\mathrm{e}^{\mathrm{i}\omega^2\frac{\beta_2}{2}(z-z_0)} \quad \text{with} \quad \tilde{\psi}_0(\omega) = \frac{1}{\sqrt{2\pi}}\int \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t}\psi_0(t). \quad [0.25 \text{ Points}] \tag{10.21}$$

Therefore, the dispersion adds a phase term to the wave. For non-monochromatic beams this may change the pulse-shape of the wave. [0.25 Points]

b) As hinted we analyse the Fourier transform

$$\tilde{\Psi}(z,k_x) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}x\, \mathrm{e}^{\mathrm{i}k_x x}\Psi(z,x). \tag{10.22}$$

The Helmholtz equation has the same mathematical structure as Eq. (10.1), therefore we can directly write down the solution

$$\tilde{\Psi}(z,k_x) = \tilde{\Psi}_0(k_x)\mathrm{e}^{-\mathrm{i}k_x^2\frac{\lambda}{4\pi}(z-z_0)} \quad \text{with} \quad \tilde{\Psi}_0(k_x) = \frac{1}{\sqrt{2\pi}}\int \mathrm{d}x\, \mathrm{e}^{\mathrm{i}k_x t}\Psi_0(x). \quad [0.25 \text{ Points}] \tag{10.23}$$

Both, dispersion and diffraction, induce a quadratic phase term to the Fourier-transformed wave. Dispersion is related to the temporal content of the wave (when it is not monochromatic) and diffraction is associated to the spatial propagation (appears when the wave is not a plane wave). In both cases, they leave the optical spectrum of the wave unchanged. Note, in order to have the same sign in the exponential term, we have to consider a dispersion coefficient β_2 which is negative. [0.25 Points]

c) Since we assume $|\psi(z,t)|^2$ to be independent of z, the solution is simply given by an exponential

$$\psi(z,t) = \mathrm{e}^{\mathrm{i}\gamma|\psi(z,t)|^2(z-z_0)}\psi_0(t) \quad \text{with} \quad \psi_0(t) = \psi(0,t). \quad [0.25 \text{ Points}] \tag{10.24}$$

With this solution our assumption holds true, i.e. $|\psi(z,t)|^2 = |\psi(0,t)|^2$. Contrary to the case of dispersion, a phase is added to the temporal domain not the spectral domain. Therefore, the phase of the propagation pulse is modulated, depending on its intensity, along its propagation direction. [0.25 Points]

d) We need to analyse a symmetric bell-shaped intensity profile up to second-order. Therefore we can write the Gaussian up to second-order in t

$$|\psi(z,t)|^2 \approx |\psi_{\max}|^2\left(1 - \frac{t^2}{K^2}\right). \tag{10.25}$$

Plugging this approximation into the solution of the previous task yields

$$\psi(d,t) = \psi_0(t)\exp\left[\mathrm{i}\gamma|\psi_{\max}|^2\left(1-\frac{t^2}{K^2}\right)z\right] = \psi_0(t)\exp\left(\mathrm{i}\gamma\psi_{\max}(d-z_0) - \mathrm{i}\gamma|\psi_{\max}|^2\frac{(d-z_0)}{K^2}t^2\right). \tag{10.26}$$

[0.5 Points]

While the first term leads to a homogeneous phase shift of the whole pulse, the second term adds a quadratic phase shift in the temporal domain, which is a temporal analog to the lens term. This quadratic phase term would lead to a temporal focusing of the propagating pulse if in addition dispersion is considered, just like a freely propagating beam, following the diffraction of the Helmholtz equation, is focused by a 1D lens in space. [0.5 Points]

e) Let us first calculate the expression for the intensity profiles of the soliton and the Peregrine breather.

Before we start, we take note of the hint and recall the series expansion of the exponential:

$$\mathrm{e}^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \dots. \tag{10.27}$$

Using the approximation up to second-order we obtain for the soliton

$$|A_s(\xi,\tau)|^2 = \left|\mathrm{sech}(\tau)\mathrm{e}^{\mathrm{i}\xi/2}\right|^2 \tag{10.28}$$

$$= \frac{4}{(\mathrm{e}^{\tau}+\mathrm{e}^{-\tau})^2} \tag{10.29}$$

$$= \frac{4}{2+\mathrm{e}^{2\tau}+\mathrm{e}^{-2\tau}} \tag{10.30}$$

$$\approx \frac{4}{2+1+2\tau+\frac{4\tau^2}{2}+1-2\tau+\frac{4\tau^2}{2}} \tag{10.31}$$

$$= \frac{1}{1+\tau^2}. \quad \text{[0.5 Points]} \tag{10.32}$$

For the Peregrin breather we obtain

$$|A_p(\xi,\tau)|^2 = \left|\left[1-\frac{4\,(1+2\mathrm{i}\xi)}{1+4\tau^2+4\xi^2}\right]\mathrm{e}^{\mathrm{i}\xi}\right|^2 \tag{10.33}$$

$$= \left|\frac{1+4\tau^2+4\xi^2-4-\mathrm{i}8\xi}{1+4\tau^2+4\xi^2}\right|^2 \tag{10.34}$$

$$= \frac{\left(1+4\tau^2+4\xi^2-4\right)^2+8^2\xi^2}{(1+4\tau^2+4\xi^2)^2} \tag{10.35}$$

$$= \frac{\left(1+4\tau^2+4\xi^2\right)^2-8\left(1+4\tau^2+4\xi^2\right)+16+8^2\xi^2}{(1+4\tau^2+4\xi^2)^2} \tag{10.36}$$

$$= 1+\frac{8-32\tau^2+32\xi^2}{(1+4\tau^2+4\xi^2)^2} \tag{10.37}$$

$$= 1+8\frac{1-4\tau^2+4\xi^2}{(1+4\tau^2+4\xi^2)^2}. \quad \text{[0.5 Points]} \tag{10.38}$$

Let's analyse these two solutions. The intensity profile of the soliton is approximated by a Lorentzian and does not depend on ξ, so the pulse intensity stays constant during propagation. The peak power rises from 0 at $\tau \to \pm\infty$ to 1 at $\tau = 0$.

The Peregrine breather in contrast does depend on the ξ-coordinate. For $\xi \to \pm\infty$ the intensity profile is just a flat line. At $\xi = 0$ it is given by

$$|A_p(0,\tau)|^2 = 1 + 8\frac{1-4\tau^2}{(1+4\tau^2)^2}. \tag{10.39}$$

This function peaks at a value of 9 at $\tau = 0$ and decays to 1 for $\tau \to \pm\infty$. The peak power at $\tau = 0$ in dependence of ξ is given by

$$|A_p(\xi,0)|^2 = 1 + \frac{8}{1+4\xi^2}, \tag{10.40}$$

which is a Lorentzian function plus a constant offset. [0.5 Points] for explanation of Peregrine breather

The soliton solution with and without the second-order approximation, the Peregrine breather solution for $\xi = 0$ and $\xi = \pm\infty$ and the peak power of both solutions in dependence of ξ are plotted in Fig. 10.2.

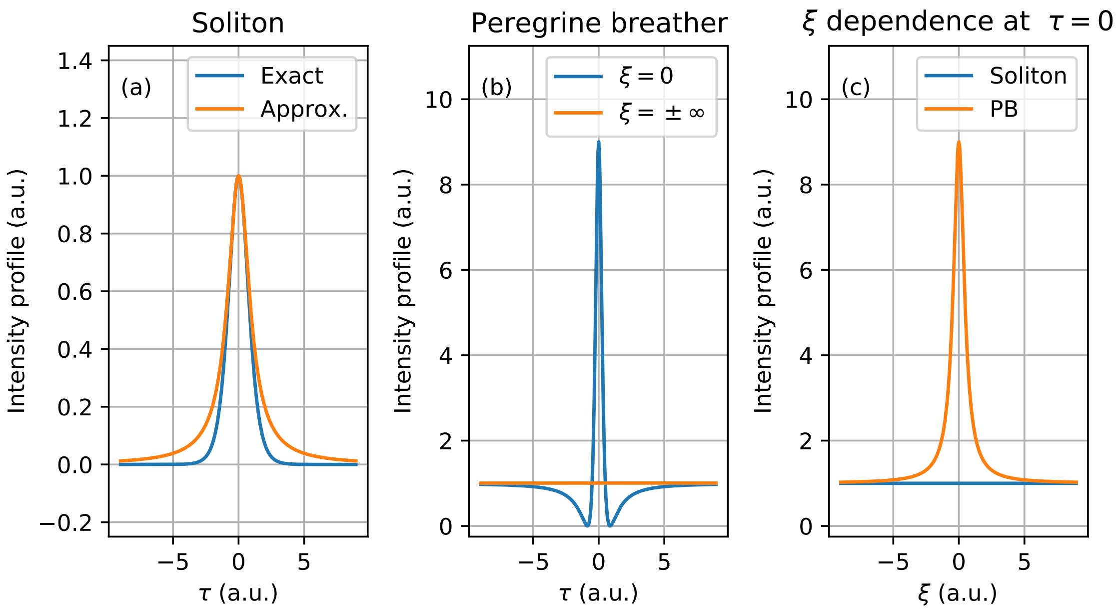


Figure 10.2 (a) Intensity profile $|A_s(\xi,\tau)|^2$ of the soliton solution with (orange) and without (blue) approximation. (b) Intensity profile $|A_P(\xi,\tau)|^2$ of the Peregrine breather solution for $\xi = 0$ (blue) and $\xi = \pm\infty$ (orange). (c) Peak power in dependence of ξ for the soliton (blue) and Peregrine breather (orange) solutions.

The interpretation of the solutions is the following: The soliton is a special solution for which the effects of dispersion and nonlinear self focusing cancel, therefore it can propagate with a constant intensity profile. The Peregrine breather in contrast does not maintain its shape, it emerges from a continuous wave and then focuses to a pulse lying over a continuous background, before it finally disappears again. [0.5 Points] for explanation of soliton

f) This exercise is just a straightforward calculation. To make it easier we calculate the different terms of the nonlinear Schrödinger equation separately.

$$\mathrm{i}\frac{\partial \psi_\varepsilon(z,t)}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \psi_\varepsilon(z,t)}{\partial t^2} - \gamma \left|\psi_\varepsilon(z,t)\right|^2 \psi_\varepsilon(z,t), \tag{10.41}$$

also keep in mind, that ψ_0 is real. The terms are

$$\mathrm{i}\frac{\partial \psi_\varepsilon(z,t)}{\partial z} = -\gamma P_0 \left[\psi_0 + \varepsilon(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right) + \frac{\partial \varepsilon(z,t)}{\partial z} \exp\left(\mathrm{i}\gamma P_0 z\right), \tag{10.42}$$

$$\frac{\beta_2}{2}\frac{\partial^2 \psi_\varepsilon(z,t)}{\partial t^2} = \frac{\beta_2}{2}\frac{\partial^2 \varepsilon(z,t)}{\partial t^2} \exp\left(\mathrm{i}\gamma P_0 z\right) \quad \text{[0.25 Points]} \tag{10.43}$$

and

$$-\gamma \left|\psi_\varepsilon(z,t)\right|^2 \psi_\varepsilon(z,t) = -\gamma \left|\psi_0 + \varepsilon(z,t)\right|^2 \left[\psi_0 + \varepsilon(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right) \tag{10.44}$$

$$\approx -\gamma \left[\psi_0^2 + \varepsilon(z,t)\psi_0 + \varepsilon^*(z,t)\psi_0\right] \left[\psi_0 + \varepsilon(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right) \tag{10.45}$$

$$\approx -\gamma P_0 \left[\psi_0 + \varepsilon(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right) - \gamma P_0 \left[\varepsilon(z,t) + \varepsilon^*(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right). \tag{10.46}$$

[0.25 Points]

We notice that $-\gamma P_0 \left[\psi_0 + \varepsilon(z,t)\right] \exp\left(\mathrm{i}\gamma P_0 z\right)$ appears on both sides of the equation and that all terms can be divided by $\exp\left(\mathrm{i}\gamma P_0 z\right)$ [0.5 Points]. Therefore, the final result is

$$\mathrm{i}\frac{\partial \varepsilon(z,t)}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \varepsilon(z,t)}{\partial t^2} - \gamma P_0 \left[\varepsilon(z,t) + \varepsilon^*(z,t)\right]. \tag{10.47}$$

g) Here, we need to insert the suggested expression for $\varepsilon(z,t)$ into Eq. (10.13), we obtain

$$-a_1 k \mathrm{e}^{\mathrm{i}(kz-\omega t)} + a_2 k^* \mathrm{e}^{-\mathrm{i}(k^* z-\omega t)} = -\omega^2 \frac{\beta_2}{2}\left[a_1 \mathrm{e}^{\mathrm{i}(kz-\omega t)} + a_2 \mathrm{e}^{-\mathrm{i}(k^* z-\omega t)}\right] \tag{10.48}$$

$$-\gamma P_0 \left[\left(a_1 + a_2^*\right) \mathrm{e}^{\mathrm{i}(kz-\omega t)} + \left(a_1^* + a_2\right) \mathrm{e}^{-\mathrm{i}(k^* z-\omega t)}\right]. \quad \text{[0.5 Points]} \tag{10.49}$$

$$\left[\left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 - k\right) a_1 + \gamma P_0 a_2^*\right] \mathrm{e}^{\mathrm{i}(kz-\omega t)} = -\left[\gamma P_0 a_1^* + \left(\frac{\beta}{2}\omega^2 + \gamma P_0 + k^*\right) a_2\right] \mathrm{e}^{-\mathrm{i}(k^* z-\omega t)} \tag{10.50}$$

This equation can only be fulfilled, if the left- and right-hand terms proportional to $\exp\left[\mathrm{i}(kz-\omega t)\right]$ and $\exp\left[\mathrm{i}\left(k^* z-\omega t\right)\right]$ fulfill their sub equation, we therefore, obtain the final set of equations

$$\left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 - k\right) a_1 + \gamma P_0 a_2^* = 0$$

$$\gamma P_0 a_1^* + \left(\frac{\beta}{2}\omega^2 + \gamma P_0 + k^*\right) a_2 = 0. \tag{10.51}$$

By complex conjugating the second equation, this set of coupled equations can be cast into matrix form

$$\begin{pmatrix} \left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 - k\right) & \gamma P_0 \\ \gamma P_0 & \left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 + k\right) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad \text{[0.5 Points]} \tag{10.52}$$

Only if the determinant is zero, there can be a non-trival solution. Therefore, we can calculate k via the determinant

$$0 = \left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 - k\right)\left(\frac{\beta_2}{2}\omega^2 + \gamma P_0 + k\right) - \gamma^2 P_0^2 \quad [0.5 \text{ Points}] \tag{10.53}$$

$$= \left(\frac{\beta_2}{2}\omega^2 + \gamma P_0\right)^2 - k^2 - \gamma^2 P_0^2 \tag{10.54}$$

$$\rightarrow k = \pm\sqrt{\frac{\beta_2^2}{4}\omega^4 + \beta_2\omega^2\gamma P_0}. \tag{10.55}$$

Therefore the final results is

$$k = \pm\frac{1}{2}|\beta_2\omega|\sqrt{\omega^2 + \frac{4\gamma P_0}{\beta_2}}. \quad [0.5 \text{ Points}] \tag{10.56}$$

h) In the normal regime of dispersion, k is always real so that no modulation instability can develop. In the anomalous regime of dispersion ($\beta_2 < 0$), modulation instability can occur when

$$\omega^2 + \frac{4\gamma P_0}{\beta_2} < 0. \tag{10.57}$$

From here we can derive a critical frequency

$$\omega_c = 2\sqrt{\frac{\gamma P_0}{|\beta_2|}}. \quad [0.5 \text{ Points}] \tag{10.58}$$

If $\omega < \omega_c$, k will be purely imaginary and there can be a modulation instability. If $\omega > \omega_c$ there can not be any instability as k is real.
In order to find the maximum gain, we calculate the maximum of k^2. It is clear that there will be a maximum, as $k = 0$ for $\omega = 0$ and $\omega = \omega_c$. Working with k^2 is easier than working with k

$$0 = \frac{\partial k^2}{\partial \omega} \tag{10.59}$$

$$0 = \beta_2^2\omega^3 + 2\beta_2\omega\gamma P_0 \tag{10.60}$$

$$= \beta_2\omega\left(\beta_2\omega^2 + 2\gamma P_0\right). \tag{10.61}$$

As $\beta_2 < 0$ the derivative is zero for $\omega_{\max} = \sqrt{\frac{2\gamma P_0}{|\beta_2|}}$. Note, that $\omega_c = \sqrt{2}\omega_{\max}$. Inserting $\omega_{\max}$ into the expression for k yields

$$k = \pm\sqrt{\frac{\beta_2^2}{4}\omega^4 + \beta_2\omega^2\gamma P_0} \tag{10.62}$$

$$= \pm\sqrt{\frac{\beta_2^2}{4}\frac{4\gamma^2 P_0^2}{\beta_2^2} - |\beta_2|\frac{2\gamma^2 P_0^2}{|\beta_2|}} \tag{10.63}$$

$$= \pm\sqrt{1-2}\gamma P_0 = \pm \mathrm{i}\gamma P_0. \quad [0.5 \text{ Points}] \tag{10.64}$$

Therefore, the maximal gain is $|k| = \gamma P_0$. The shape of the gain needs to be the following. It is zero at $\omega = 0$ and $\omega = \omega_c$. For small ω it needs to rise almost linearly, as the ω^2 term in the square root of Eq. (10.56) can be neglected. At the maximum at ω_{max} the parabola in the square root takes over and the gain drops relatively quickly to zero. The gain function is plotted in Fig. 10.3. [0.5 Points]

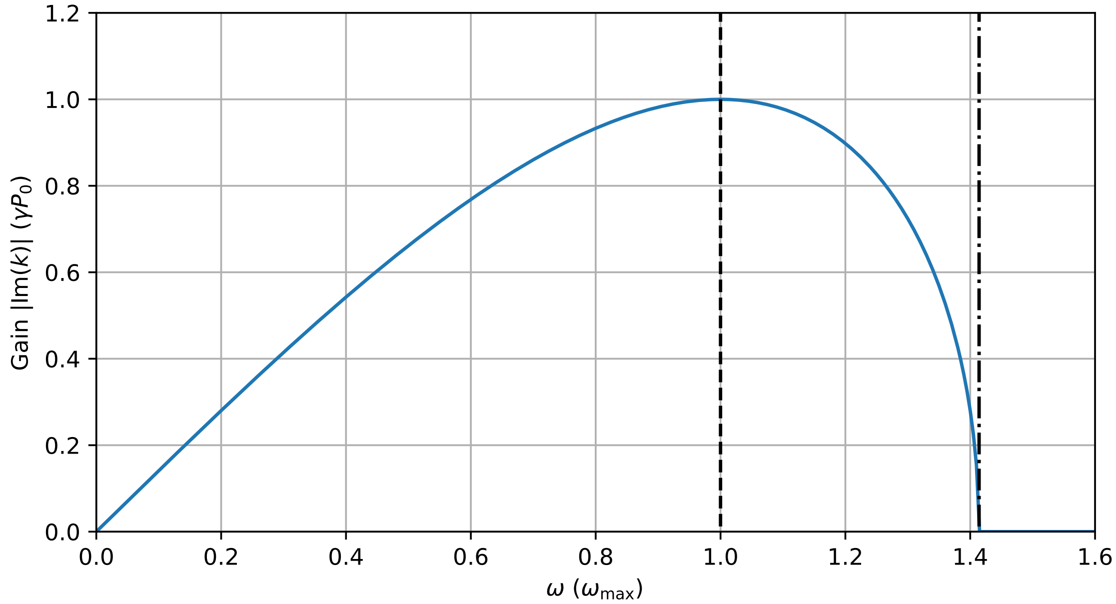


Figure 10.3 Gain profile ($|\text{Im}(k)|$) of the modulation instability. The dashed line shows the position of maximal gain $\omega_{\text{max}} = \sqrt{\frac{2\gamma P_0}{|\beta_2|}}$. The dashed-dotted line shows the position of the critical frequency $\omega_c = \sqrt{2}\omega_{\text{max}}$ above which the gain is zero.

i) The mathematical form of Eq. (10.17) is identical to the normalized nonlinear Schrödinger equation (10.7). Therefore, the two equations share the same solutions. As we have seen the Peregrine breather solution forms out of a homogeneous background and peaks sharply at a certain position. Consequently, the solution of the Peregrine breather is clearly a prototype of the so-called rogue wave. [1 Point]