Hwk #30:

Use the multi-variable chain rule to determine $f'(x)$, when $f(x) := \int_0^x \frac{\sin(xt)}{t}\,dt$.

Analogous question for $g(x) := \int_{x/2}^{2x} \frac{e^{xt}}{t}\,dt$.

Again, we rely on the Math 447 expert, who tells us that it is legitimate to move derivatives past the integral sign in this example.

Solution: This time, x occurs in two places in the formula for $f(x)$. The MV chain rule, written in terms of partials, tells us to consider each location separately and apply a partial (single variable) derivative (as in the single variable chain rule), and then to add the results obtained for each separate location. More formally, we consider

$$F(u,v) := \int_0^u \frac{\sin vt}{t}\,dt$$

and we substitute $u = x$ and $v = x$ to get $f(x) = F(x,x)$. So we have

$$f'(x) = (\partial_1 F)(x,x)\frac{\partial u}{\partial x} + (\partial_2 F)(x,x)\frac{\partial v}{\partial x}$$

For $\partial_1 F$, we use the fundamental theorem (derivative of an antiderivative) to get $(\partial_1 F)(u,v) = \frac{\sin vu}{u}$. For $\partial_2 F$, we use differentiation under the integral sign (with permission from the M447 expert again given specifically for the situation of this problem, not as a blank cheque!) and get $(\partial_2 F)(u,v) = \int_0^u \frac{t\cos vt}{t}\,dt = [\frac{1}{v}\sin vt]_{t=0}^{t=u} = \frac{\sin vu}{v}$. Putting it all together (with $\partial u/\partial x = 1 = \partial v/\partial x$ because $u = x$ and $v = x$), we get

$$f'(x) = \frac{\sin(x^2)}{x} + \frac{\sin(x^2)}{x} = 2\frac{\sin(x^2)}{x}\ .$$

The same works for g, in principle: We define $G(u,v,w) := \int_w^u \frac{e^{vt}}{t}\,dt$ and let $w = x/2$, $u = 2x$, and $v = x$. So $g(x) = G(2x, x, \frac{x}{2})$. Note that derivatives with respect to the *lower* limit of integration get a minus sign from the fundamental theorem, and that we have inner derivatives $\frac{\partial w}{\partial x} = \frac{1}{2}$ and $\frac{\partial u}{\partial x} = 2$ this time.

$$g'(x) = 2\frac{e^{x\cdot 2x}}{2x} + \int_{x/2}^{2x} \frac{te^{xt}}{t}\,dt - \frac{1}{2}\frac{e^{x\cdot x/2}}{x/2} = \frac{e^{2x^2}}{x} + \left[\frac{e^{xt}}{x}\right]_{t=x/2}^{t=2x} - \frac{e^{x^2/2}}{x} = \frac{2e^{2x^2}}{x} - \frac{2e^{x^2/2}}{x}$$

Hwk #31:

A quantity w depends on the coordinates x, y, z in 3-space as follows: $w = x^2 + y^2 + xyz$ (1). We study w especially on the plane given by $z = x + 2y$. Then we have there $w = x^2 + y^2 + xy(x+2y) = x^2 + y^2 + x^2y + 2xy^2$ (2).

Now we calculate $\frac{\partial w}{\partial x}$ from (1): $\frac{\partial w}{\partial x} = 2x + yz$. On the plane, this simplifies to $\frac{\partial w}{\partial x} = 2x + y(x+2y) = 2x + xy + 2y^2$.

Calculating $\frac{\partial w}{\partial x}$ on the plane directly from (2), we get $\frac{\partial w}{\partial x} = 2x + y(x+2y) = 2x + 2xy + 2y^2$. We clearly have a discrepancy by a term xy. What is wrong? Clear up the confusion. (This requires some text as well as formulas.)

Solution: When we calculate $\frac{\partial w}{\partial x}$ from (1), we are taking a partial derivative with respect to z of a three-variable function f given by $f(x,y,z) = x^2 + y^2 + xyz$. In this partial derivative, both y and z are treated as constant. The partial derivative $\frac{\partial w}{\partial x}$ is a directional derivative in direction $[1\,,\ 0\,,\ 0]^T$, which is a direction that goes off the plane $z = x + 2y$, even if we later evaluate this derivative in a point on that plane.

In contrast, when we substitute first $z = x + 2y$ and then take the partial derivative $\frac{\partial w}{\partial x}$ form (2), we are taking a partial derivative of a two-variable function g given by $g(x,y) = x^2 + y^2 + xy(x+2y)$. In this partial derivative, y is still constant and x still varies, as before; but z is now not constant, but also varies because $z = x + 2y$. For f, this is now a directional derivative in direction $[1\,,\ 0\,,\ 1]^T$, which is a direction lying *in* the plane $z = x + 2y$.

The notation $\frac{\partial w}{\partial x}$ is ambiguous in this context, because the variable w can refer to two different functions f versus g. Unless the ambiguity is resolved by an explaining text or context, the notation $\frac{\partial w}{\partial x}$ should be avoided in such a situation. Rather the result according to part (1) should be written as $\partial_1 f(x, y, x+2y)$, and the result at part (2) should be written as $\partial_1 g(x,y)$.

We then have

$$(*)\qquad \begin{aligned} \frac{\partial}{\partial x} f(x,y,x+2y) &= \partial_1 f(x,y,x+2y)\frac{\partial x}{\partial x} + \partial_2 f(x,y,x+2y)\frac{\partial y}{\partial x} + \partial_3 f(x,y,x+2y)\frac{\partial (x+2y)}{\partial x} \\ &= \partial_1 f(x,y,x+2y) + \partial_3 f(x,y,x+2y) \end{aligned}$$

and

$$(**)\qquad \frac{\partial}{\partial x} g(x,y) = \partial_1 g(x,y)$$

Both of these results are equal and correspond to calculation (2), whereas calculation (1) gave just the first term of the sum arising in $(*)$.

Hwk #32:

This example is taken from *I. Rosenholtz, L. Smylie: "The only Critical Point in Town" Test*, Mathematics Magazine **58**(1985), 149–150.

Show that the function

$$g : (x,y) \mapsto y^2 + 3(y + e^x - 1)^2 + 2(y + e^x - 1)^3\,,\ \mathbb{R}^2 \to \mathbb{R}$$

has exactly one critical point, and that this point is a relative mimimum.

Furthermore explain why this point is NOT an absolute minimum.

Solution: It is convenient to use the abbreviation $u := y + e^x - 1$.

$$\frac{\partial g(x,y)}{\partial x} = 6ue^x + 6u^2e^x \qquad \frac{\partial g(x,y)}{\partial y} = 2y + 6u + 6u^2$$

For both to vanish we need $y = 0$ and $u + u^2 = 0$, which latter means $u = 0$ or $u = -1$. Since $y = 0$, $u = e^x - 1$, and this cannot equal -1. So $u = 0$, and this means $x = 0$.

Therefore the only critical point is $(x,y) = (0,0)$. Let's calculate the Hessian:

$$Hg(x,y) = \begin{bmatrix} (6u + 6u^2)e^x + (6 + 12u)e^{2x} & (6+12u)e^x \\ (6+12u)e^x & 8 + 12u \end{bmatrix}$$

$$Hg(0,0) = \begin{bmatrix} 6 & 6 \\ 6 & 8 \end{bmatrix}$$

This matrix is positive definite, by the Hurwitz test: $h_{11} = 6 > 0$ and $h_{11}h_{22} - h_{12}^2 = 12 > 0$. Therefore the origin is a relative minimum of g. $g(0,0) = 0$. But $g(0,y) = 4y^2 + 2y^3 \to -\infty$ as $y \to -\infty$. The function is unbounded below and does not have an absolute minimum.

Comment: This example shows two things. Firstly, it does not suffice to check the values of a function at the only candidates for a relative minimum in order to determine a global minimum, *unless* the existence of an absolute minimum is established beforehand.

But there is a second lesson hidden in this example. An intuitively plausible argument would go like this: "If I am hiking in a landskape and standing in a local minimum, but there are points out there with lower elevation than my present location, then it must be possible, in principle, to reach them by a walk that goes through some pass (mathematically a saddle point). So, if a local minimum is not a global minimum, there should be another critical point, a saddle point, somewhere." This argument does actually have some merit. However, the saddle point could have run away to infinity. In our example the 2-variable function $(u,y) \mapsto y^2 + 3u^2 + 2u^3$ has a local minimum at $(u,y) = (0,0)$ and a saddle point at $(u,y) = (-1,0)$. But $u \to -1$ corresponds to $x \to -\infty$ when $u = y + e^x - 1$ and $y = 0$.

In advanced applications, this principle 'there should be a saddle point, if only we can make sure that it hasn't run off to infinity' is a powerful tool in solving partial differential equations.

Hwk #33:

> This example is taken from Marsden-Tromba: Show that the function f given by $f(x,y) = (y - 3x^2)(y - x^2)$ has a critical point in the origin, which is neither a relative minimum nor a relative maximum. What kind of '***'definite is the Hessian?
>
> Show also that all single-variable radial functions $t \mapsto f(t\cos\phi, t\sin\phi)$ have a relative minimum at $t = 0$.

Solution: For fixed ϕ, let $g(t) := f(t\cos\phi, t\sin\phi) = t^2\sin^2\phi - 4t^3\sin\phi\cos^2\phi + 3t^4\cos^4\phi$. Then $g'(0) = 0$ and $g''(0) = 2\sin^2\phi$. For $\sin\phi \neq 0$, we can argue that g has a local minimum at 0 because the first derivative vanishes and the second derivative is positive there. For $\sin\phi = 0$, we have $\cos^2\phi = 1$ and $g(t) = 3t^4$, and we again have a local minimum at 0, albeit a 'degenerate' one that cannot be detected by the second derivative test. .

Now clearly $f(0,0) = 0$, but there are both positive and negative values in any neighbourhood of $(0,0)$. For instance $f(x, 2x^2) = -x^4 < 0$. Let's calculate the Hessian from $f(x,y) = 3x^4 - 4x^2y + y^2$: $f_{xx}(x,y) = 36x^2 - 8y$, $f_{xy}(x,y) = -8x$, $f_{yy}(x,y) = 2$. So $Hf(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$. This matrix is positive semidefinite, but not positive definite.

Note: Similar as with continuity and differentiability, we learn here for the local minimu property that the MV version cannot be captured by having the single variable version in all directions. However, a slightly strengthened version of the single variable local minimality (namely that the sufficient condition of having positive second derivative is satisfied) in all directions does suffice to prove local minimality in the multi-variable context for C^2 functions.

Hwk #34:

> *This example is geometrically appealing, but alas calculationally lengthy. This is why I give you the intermediate steps and hints to navigate you through. Ideally, it should be*

done with the help of symbolic algebra software, and you are welcome to use this tool, if available.

We want to find a shortest connection between two plane curves, namely $y = x^2 + 2$ and $y = \frac{1}{2}(x-1)^2$. A precise plot is attached. Choose points $P = (a, a^2+2)$ on the first parabola and $Q = (b, \frac{1}{2}(b-1)^2)$ on the second and minimize the square of the distance. Determine all critical points and classify them. Does one of them probide a *global* minimum? Why?

Hint 1: While it is possible to take one of the equations for a critical point and solve it for b via by means of the quadratic formula, and then plug in the result in the other equation, this is tedious. It is more straightforward to take successively linear combinations of the two equations with the strategy of first eliminating b^3, then b^2, then b, until one polynomial equation in a remains.

Hint 2: After an obvious factorization of this polynomial equation, an easy solution $a = 1$ can be guessed, and when this is factored off, a 4th order polynomial remains that can be factored into two quadratics with integer coefficients; indeed one factor is $a^2 + 2a + 3$.

Solution: We take the distance squared between a point $P = (a, a^2 + 2)$ on the first parabola and a point $Q = (b, \frac{1}{2}(b-1)^2)$ on the second parabola and get the function

$$f(a,b) := (a-b)^2 + \Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big)^2 .$$

We obtain

$$\partial f(a,b)/\partial a = 2(a-b) + 4a\Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big) = 0$$
$$\partial f(a,b)/\partial b = 2(b-a) - 2(b-1)\Big(a^2 + 2 - \frac{1}{2}(b-1)^2\Big) = 0$$

So we have to solve the system of two polynomial equations

$$-b + 2a^3 + 4a - ab^2 + 2ab = 0 \quad (1)$$
$$2a^2 - 2a + 3 - 2a^2 b + b^3 - 3b^2 + b = 0 \quad (2)$$

Following the hint, we take a times (2) plus b times (1) to get rid of b^3. We obtain

$$2a^3 - 2a^2 + 3a + 5ab - (a+1)b^2 = 0 \quad (3)$$

Now we add $-(a+1)$ times (1) to a times (3) to eliminate b^2. We obtain

$$-4a^3 - a^2 - 4a + b(3a^2 - a + 1) = 0 \quad (4)$$

We solve this for b and plug it back into (1):

$$2a^3 + 4a + (2a-1)\frac{4a^3 + a^2 + 4a}{3a^2 - a + 1} - a\left(\frac{4a^3 + a^2 + 4a}{3a^2 - a + 1}\right)^2 = 0$$

Clearing denominators and expanding, we get

$$2a^7 + 4a^6 + 3a^5 - 5a^4 - 7a^3 + 3a^2 = 0 \quad (5)$$

We can factor off $a = 0$ twice and guess the solution $a = 1$, thus factoring off $(a-1)$. After long division of polynomials, we get

$$0 = 2a^4 + 6a^3 + 9a^2 + 4a - 3 = (2a^2 + 2a - 1)(a^2 + 2a + 3)$$

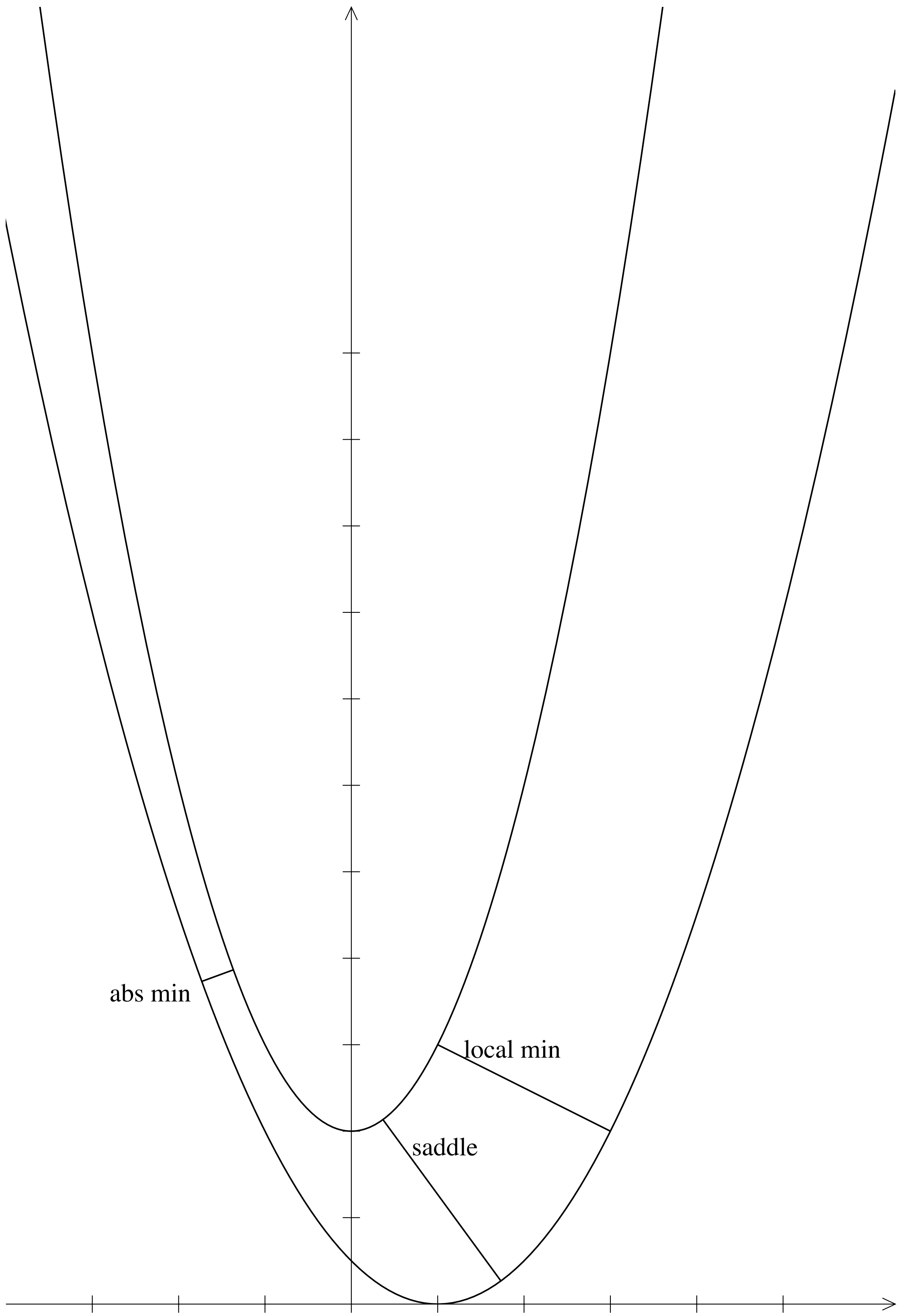

Figure 1: Figure for Hwk # 34

This factorization, short of the hint given, would best be found by means of a symbolic algebra software. $a^2+2a+3=0$ doesn't have real solutions; $2a^2+2a-1=0$ can be solved by the quadratic formula. We have thus found the following solutions of (5):

$$a_0 = 0\,, \quad a_1 = 1\,, \quad a_{2,3} = \frac{1}{2}\Big(-1 \pm \sqrt{3}\Big)$$

We can get b form a by means of (4), but initially, all of these are mere *consequences* of (1),(2); they are not *equivalent* to (1),(2). We must plug them back into the original equations. Clearly $a_0 = 0$ violates (2). For $a \neq 0$, (1),(2) $\iff$ (1),(3) $\iff$ (1),(4) $\iff$ (4),(5). The last equivalence uses $3a^2 - a + 1 \neq 0$. So we have found

$$\begin{array}{lll}
(a_1, b_1) = (1,3) & P_1 = (1,3) & Q_1 = (3,2) \\[1ex]
(a_2, b_2) = \left(\frac{1}{2}\Big(-1+\sqrt{3}\Big), \sqrt{3}\right) & P_2 = \left(\frac{1}{2}\Big(-1+\sqrt{3}\Big), 3 - \frac{1}{2}\sqrt{3}\right) & Q_2 = \Big(-\sqrt{3}, 2-\sqrt{3}\Big) \\[1ex]
(a_3, b_3) = \left(\frac{1}{2}\Big(-1-\sqrt{3}\Big), -\sqrt{3}\right) & P_3 = \left(\frac{1}{2}\Big(-1-\sqrt{3}\Big), 3 + \frac{1}{2}\sqrt{3}\right) & Q_3 = \Big(\sqrt{3}, 2+\sqrt{3}\Big)
\end{array}$$

Numerical values are:

$$\begin{array}{ll|l}
P_1 = (1,3) & Q_1 = (3,2) & |P_1Q_1| = 2.236 \\[1ex]
P_2 = (0.366, 2.134) & Q_2 = (1.732, 0.268) & |P_2Q_2| = 2.313 \\[1ex]
P_3 = (-1.366, 3.866) & Q_3 = (-1.732, 3.732) & |P_3Q_3| = 0.390
\end{array}$$

The Hessian of f is

$$Hf(a,b) = \begin{bmatrix} 8+12a^2+4b-2b^2 & -2+4a-4ab \\ -2+4a-4ab & 1-2a^2-6b+3b^2 \end{bmatrix}$$

$$Hf(a_1,b_1) = \begin{bmatrix} 14 & -10 \\ -10 & 8 \end{bmatrix} \qquad \text{pos def}$$

$$Hf(a_2,b_2) = \begin{bmatrix} 14-2\sqrt{3} & -10+4\sqrt{3} \\ -10+4\sqrt{3} & 8-5\sqrt{3} \end{bmatrix} \approx \begin{bmatrix} 10.536 & -3.072 \\ -3.072 & -0.660 \end{bmatrix} \qquad \text{indefinite}$$

$$Hf(a_3,b_3) = \begin{bmatrix} 14+2\sqrt{3} & -10-4\sqrt{3} \\ -10-4\sqrt{3} & 8+5\sqrt{3} \end{bmatrix} \approx \begin{bmatrix} 17.464 & -16.928 \\ -16.928 & 16.660 \end{bmatrix} \qquad \text{pos def}$$

So (a_2, b_2) is a saddle point, the other two are local minima. The one with the smaller distance, namely (a_3, b_3) can be accepted as a global minimum, PROVIDED we know apriori that a global minimum EXISTS.

We cannot argue directly, because $\mathbb{R} \times \mathbb{R} \ni (a,b)$ is not bounded. However, we have the additional feature that $f(a,b) \to \infty$ as either a or b goes to infinity. For instance, when asking whether a global minimum exists, we can a-priori neglect all (a,b) for which $f(a,b) > 100$, because at some points f has smaller values, e.g., $f(0,1) = 5 \ll 100$. So we only need to consider those (a,b) where $|a-b| \leq 10$, because otherwise $f(a,b) > 10^2 + 0$. But if $|a-b| \leq 10$, we have $a^2 + 2 - \frac{1}{2}(b-1)^2 = \frac{1}{2}a^2 + \frac{1}{2}[a^2 - (b-1)^2] + 2 \geq \frac{1}{2}a^2 + 2 - \frac{1}{2}|a-b+1|\,|a+b-1| \geq \frac{1}{2}a^2 + 2 - \frac{11}{2}(2|a|+11)$. For $|a|$ sufficiently large, this will again be > 10, making $f(a,b) > 0 + 10^2$. So we need to consider only the absolute minimum on some closed and bounded set $\{(a,b) : |a-b| \leq 10\,,\ |a| \leq C\}$, and an abolute minimum exists on this set. Outside ths set, the values of f are larger, so we do have an absolute minimum on $\mathbb{R}^2$.

Hwk #35:

Suppose in the following matrices, the starred entries are not known. Which of the five possibilities 'positive definite', 'positive semidefinite (but not definite)', 'negative definite', 'negative semidefinite (but not definite)', 'indefinite' remains a possibility, based on knowledge only of the known entries?

$$(a)\begin{bmatrix} 3 & * \\ * & * \end{bmatrix} \qquad (b)\begin{bmatrix} * & * \\ * & -5 \end{bmatrix} \qquad (c)\begin{bmatrix} * & 6 \\ 6 & * \end{bmatrix} \qquad (d)\begin{bmatrix} 3 & * \\ * & -1 \end{bmatrix} \qquad (e)\begin{bmatrix} 0 & 1 & * \\ 1 & * & * \\ * & * & * \end{bmatrix}$$

Consider the following: While the Hurwitz test was worded in a way to calculate determinants starting from the top, the order in which the variables are listed (and thus determine entries of the matrix) is not essential for definiteness of a matrix; so you could use the determinants in the Hurwitz test starting at any diagonal element and then calculating 2×2, 3×3, etc. determinants, adding any one variable (row and column) at a time.

Solution: (a) this matrix could be positive (semi-)definite or indefinite, but not negative (semi-) definite, because the vector $[1\,,\ 0]^T$ makes the quadratic form positive.

(b) this matrix could be negative (semi-)definite or indefinite, but not positive (semi-)definite, because the vector $[0\,,\ -1]^T$ makes the quadratic form negative.

(c) no conclusion can be drawn: If one diagonal element is positive and one negative, then the matrix is indefinite. If both diagonal elements are larger than 6, the matrix is positive definite by Gershgorin's test. If both diagonal elements are below -6, the matrix is negative definite. If both diagonal elements are equal 6 (or equal -6), the matrix is positive (or negative) semidefinite by explicit writing down of the quadratic form. (This does not exhaust all possibilities that the $*$ could stand for, but it already represents all cases for definiteness properties of the matrix.)

(d) this matrix is indefinite, regardless of the off-diagonal elements.

(e) this matrix is indefinite. The determinant of the 2×2 submatrix $\begin{bmatrix} 0 & 1 \\ 1 & * \end{bmatrix}$ is -1, so the matrix cannot be positive definite (and the 0 in the corner also confirms this). Similarly, the negative of this matrix cannot be positive definite either. – We do not have a Hurwitz test for semi-definiteness, so we cannot rule this case out by determinants. But let's just calculate the quadratic form with the vector $[1\,,\ t\,,\ 0]^T$. We get $0\cdot 1^2 + 2\cdot 1\cdot t + *\cdot t^2 = 2t + *t^2$. Whatever the value of $*$, this expression is positive for t positive and sufficiently small, but is negative for t negative of sufficiently small absolute value.

Hwk #36:

Find the absolute minimum and absolute maximum of $x^2 + (y-1)^2 + z^2 - xyz$ on the ball $x^2+y^2+z^2 \le 3^2$. *Hint: For the boundary consideration, use the xz plane as equator plane for the spherical coordinates, to benefit from the symmetry of the problem. Otherwise formulas get obnoxiously messy.*

Solution: Since this is a continuous function on a bounded and closed set, we know that an absolute minimum and maximum *exist*. We do not need to calculate Hessians to check for local minima or maxima, because they are not asked. The absolute extrema can be found by selecting from critical points in the interior and critical points on the boundary.

Let's first study any interior critical points of the function f given by $f(x,y,z) := x^2+(y-1)^2+z^2-xyz$: For the gradient to vanish, we need

$$\begin{aligned}(\partial_1 f)(x,y,z) &= 2x - yz = 0\\ (\partial_2 f)(x,y,z) &= 2(y-1) - xz = 0\\ (\partial_3 f)(x,y,z) &= 2z - xy = 0\end{aligned}$$

Combining the first and third condition, we get $2x^2 = 2z^2 = xyz$. So we conclude either (1) $x = z$, $y = 2$ or (2) $x = -z$, $y = -2$, or (3) $x = z = 0$. Plugging each case into the 2nd equation we get: two solutions $(x,y,z) = (\pm\sqrt{2}, 2, \pm\sqrt{2})$ from (1). Two solutions $(x,y,z) = (\pm\sqrt{6}, -2, \mp\sqrt{6})$ from (2). One solution $(x,y,z) = (0,1,0)$ from (3). Since we have only been drawing conclusions from this system, we need to check that all these five solutions do satisfy all three equations. They do. The solutions from (2) are not in the ball and may therefore be discarded.

We note $f(\pm\sqrt{2}, 2, \pm\sqrt{2}) = 1$ and $f(0,1,0) = 0$.

We parametrize the boundary by spherical coordinates: $y = 3\cos\vartheta$, $x = 3\sin\vartheta\,\cos\varphi$, $z = 3\sin\vartheta\,\sin\varphi$. While this is originally meant for $\varphi \in [0,2\pi]$ and $\vartheta \in [0,\pi]$, we can extend the coordinates by 'wrapping over' to $\varphi, \vartheta \in \mathbb{R}$. This reflects the fact that the sphere does not have a boundary. We only need to look for interior critical points of the 2-variable function

$$\begin{aligned}g(\vartheta,\varphi) &:= f(3\sin\vartheta\,\cos\varphi, 3\cos\vartheta, 3\sin\vartheta\,\sin\varphi)\\ &= 9\sin^2\vartheta + (3\cos\vartheta - 1)^2 - 27\cos\vartheta\,\sin^2\vartheta\,\sin\varphi\,\cos\varphi\\ &= 10 - 6\cos\vartheta - \tfrac{27}{2}\cos\vartheta\,\sin^2\vartheta\,\sin 2\varphi\end{aligned}$$

The vanishing of g_φ requires $\cos\vartheta\,\sin^2\vartheta\,\cos(2\varphi) = 0$. This means $\vartheta \in \{0, \frac{\pi}{2}, \pi\}$ or $\varphi \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$.

The vanishing of g_ϑ requires $6\sin\vartheta + \frac{27}{2}\sin^3\vartheta\,\sin 2\varphi - 27\sin\vartheta\,\cos^2\vartheta\,\sin 2\varphi = 0$.

$\vartheta \in \{0,\pi\}$ satisfy this condition and give rise to boundary critical points $f(0,3,0) = 4$ and $f(0,-3,0) = 16$.

$\vartheta = \frac{\pi}{2}$ (and hence $y = 0$) still needs $\sin 2\varphi = -\frac{4}{9}$ for g_ϑ to vanish. We do not need to pursue these critical points further because $g(\frac{\pi}{2}, \varphi) \equiv 10$, so they would be neither absolute minima nor absolute maxima. *Note: Lest this seem confusing: the function g is constant on the 'equator' $y = 0$, $x^2 + z^2 = 9$. But not all points on the equator are critical points, because criticality also depends on how the function changes off the equator.*

$\varphi \in \{\frac{\pi}{4}, \frac{5\pi}{4}\}$ still needs $6\sin\vartheta + \frac{27}{2}\sin^3\vartheta - 27\sin\vartheta\,\cos^2\vartheta = 0$ for g_ϑ to vanish. Apart from retrieving $\vartheta \in \{0,\pi\}$, which has been discussed already, we have to solve $6 + \frac{27}{2}(1-\cos^2\vartheta) - 27\cos^2\vartheta = 0$, or $\cos\vartheta = \pm(\frac{13}{27})^{1/2}$. Then $\sin^2\vartheta = \frac{14}{27}$ and $g = 10 \mp 13(\frac{13}{27})^{1/2}$.

Similarly, $\varphi \in \{\frac{3\pi}{4}, \frac{7\pi}{4}\}$ still needs $6\sin\vartheta - \frac{27}{2}\sin^3\vartheta + 27\sin\vartheta\,\cos^2\vartheta = 0$ for g_ϑ to vanish. Apart from retrieving old critical points, we have to solve $6 - \frac{27}{2}(1-\cos^2\vartheta) + 27\cos^2\vartheta = 0$, or $\cos\vartheta = \pm(\frac{5}{27})^{1/2}$. Then $\sin^2\vartheta = \frac{22}{27}$ and $g = 10 \pm 5(\frac{5}{27})^{1/2}$, values that do not qualify for a global extremum.

Concludingly, we have found the global minimum 0 at $(x,y,z) = (0,1,0)$ and the global maximum $10 + 13(\frac{13}{27})^{1/2} \approx 19.0206$ on the boundary at $(x,y,z) = (\pm(\frac{7}{3})^{1/2}, -(\frac{13}{3})^{1/2}, \pm(\frac{7}{3})^{1/2})$.

Hwk #37:

> Redo the boundary part of the calculations from the previous problem using Lagrange multipliers.

Solution: Minimizing/maximizing the expression $f(x, y, z) := x^2 + (y-1)^2 + z^2 - xyz$ on the level set $h(x, y, z) = x^2 + y^2 + z^2 - 9 = 0$ leads to the following necessary conditions for a local extremum:

$$\begin{aligned} 2x - yz &= 2\lambda x \quad (1)\\ 2(y-1) - xz &= 2\lambda y \quad (2)\\ 2z - xy &= 2\lambda z \quad (3)\\ x^2 + y^2 + z^2 &= 9 \quad (4) \end{aligned}$$

Guided by the symmetry between x and z, we seek to simplify this system by calculating $x\cdot(1)-z\cdot(3)$. After slight rearrangement, this gives $2(1-\lambda)(x^2 - z^2) = 0$. So either $x = \pm z$ or $\lambda = 1$. We also subtract $(1) - (3)$ directly and get $(2 + y)(x - z) = 2\lambda(x - z)$. So if $x \neq z$, we need $y = 2\lambda - 2$.

The case $\lambda = 1$ leads to $yz = 0 = xy$, $xz = -2$. So since x, z cannot vanish, we need $y = 0$. The combination $xz = -2$, $x^2 + z^2 = 9$ gives four solutions (the ones on the equator in the previous problem), and for all of them, the value of f is 10. We still have to consider the cases $x = z$ and $x = -z$.

Let's now look at the case $x = z$. Either $x = z = 0$ and hence $y = \pm 3$ with values 4 or 16 for f, or else we can cancel x from (1) and get $2\lambda = 2 - y$. Plugging this into (2) we get $2(y-1) - x^2 = (2-y)y$. And (4) becomes $2x^2 + y^2 = 9$. These togehter are equivalent to $x^2 = \frac{7}{3}$, $y^2 = \frac{13}{3}$.

In the case $x = -z$ we may assume $x \neq z$ (hence $x, z \neq 0$), b/c $x = z$ has been studied already. So we have then $2\lambda = y + 2$ and this case leads to $-2 + x^2 = y^2$ from (2) and $2x^2 + y^2 = 9$ from (4). Hence $x^2 = \frac{11}{3}$ and $y^2 = \frac{5}{3}$. The values of f corresponding to this case are $10 \pm \frac{5}{3}(\frac{5}{3})^{1/2}$.

Hwk #38:

> This problem gives you the most celebrated use of Lagrange multipliers, but it requires some intrduction to appreciate it. (The calculations aren't bad at all.)
>
> A famous task in linear algebra and matrix theory is to find *eigenvalues* of a given matrix. If A is a square matrix and you can find a *non-zero* vector v such that Av is actually a multiple of v, i.e., λv where λ is a number, then we call λ an eigenvalue of the matrix A (and v an eigenvector). For instance, the matrix $A = \begin{bmatrix} 3 & 2 \\ -3 & -4 \end{bmatrix}$ has 2 as an eigenvalue and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ as a corresponding eigenvector, because
>
> $$\begin{bmatrix} 3 & 2 \\ -3 & -4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$
>
> It also has -3 as an eigenvalue with $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ as an eigenvector. Of course multiples of eignevectors are again eigenvectors, e.g., if $Av = 2v$ then also $A(7v) = 2(7v)$. — In the example, there are only these two numbers $\lambda_1 = 2$ and $\lambda_2 = -3$ that are eigenvalues. If you try to find $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ solving $Av = \lambda v$ for any other λ you will only get the solution $v_1 = v_2 = 0$, i.e., only the zero vector. (Try it, just to gain familiarity with the notions.)

This problem is about eigenvalues of *symmetric* matrices. They play a role in studying definiteness of symmetric matrices. In physics, they are key concepts in describing rotating motions of rigid bodies. To every body, there is associated a symmetric 3×3 matrix called its 'tensor of inertia', whose eigenvectors point in the directions of such axes about which the body can rotate without wobbling (i.e., in a balanced way). The eigenvalues are called the moments of inertia about these axes.

To every symmetric $n \times n$ matrix A we associate the quadratic function $f(x) := x^T Ax$ where $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$. We try to minimize or maximize $f(x)$ under the constraint $x^T x = 1$ (i.e., for x on the unit sphere).

(a) Write out $f(x)$ in components x_i for a 3×3 matrix A whose entries are called a_{ij}. Explain why a global maximum and a lobal minimum of $f(x)$ on the sphere are a-priori guaranteed to exist.

(b) Use the Lagrange multiplier method to set up equations satisfied by the x providing a minimum or a maximum. (You may have written all these in components; but now make sure to rewrite the whole stuff again in matrix and vector form.) While you are not asked to actually solve for x (that would be very tedious, involving a cubic equation for λ), I ask you to express the value of f at the minimum and maximum in terms of the Lagrange multiplier. [Be aware that when finding the max vs the min, x and λ will typically refer to numerically different quantities in these two cases.]

You have just proved that every symmetric 3×3 matrix has (at least) two real eigenvalues. (Actually, if A is a multiple of the identity matrix, these two eigenvalues coincide.) And with jut a bit more writing, the same can be done for symmetric $n \times n$ matrices.

The method can be cranked up a bit, by throwing in further constraints, to prove that every symmetric $n \times n$ matrix has n real eigenvalues (some of which may coincide). This may well be among the most important pieces of insight in undergraduate mathematics, and it's a pity that it often falls between the cracks of separating Calc 3 and Linear Algebra into independent courses of the curriculum.

(c) Show, in a very brief calculation: If A is positive definite, then all its eigenvalues are positive. If A is positve semidefinte, then all of its eigenvalues are ≥ 0. FYI: The converse is also true; so indeed a symmetric matrix is positve definite (resp. emidefinite) IF AND ONLY IF all of its eigenvalues are positive (resp. non-negative). This statement is sponsored by the above proof (a), (b) and some extra dose of linear algebra. It is the launch pad for proving the Hurwitz and Gershgorin tests I gave you before.

Solution: (a) Using the symmetry of the matrix already, we get

$$f(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

Being a polynomial, this function is in particular continuous, and we constrain it to the sphere $x_1^2 + x_2^2 + x_3^2 = 1$, which is a closed and bounded set. Therefore a minimum and a maximum of f on the sphere exist.

(b) The Lagrange multiplier method says that at a constrained minimum, and at a constrained maximum, there exists a $\lambda \in \mathbb{R}$ such that $\nabla f = \lambda \nabla g$ (where $g(x) := x^T x - 1 = 0$ is the constraint), provided ∇g does not vanish on the set $g = 0$. We can easily see $\nabla g(x) = 2x$ (and this clearly does not vanish on the constraining set $x^T x = 1$). Let's calculate $\nabla f(x)$:

$$\nabla f(x) = \begin{bmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \partial_3 f(x) \end{bmatrix} = \begin{bmatrix} 2a_{11}x_1 + 2a_{12}x_2 + 2a_{13}x_3 \\ 2a_{12}x_1 + 2a_{22}x_2 + 2a_{23}x_3 \\ 2a_{13}x_1 + 2a_{23}x_2 + 2a_{33}x_3 \end{bmatrix} = 2Ax$$

We conclude there exist $\lambda \in \mathbb{R}$ and a nonzero vector $x \in \mathbb{R}^3$ such that $2Ax = 2\lambda x$, hence λ i an eigenvalue and x is an eigenvector. (We know more about x than that it is non-zero; actually $\|x\| = 1$.)

Since $Ax = \lambda x$ at the extremum, we conclude $x^T Ax = \lambda x^T x = \lambda$, so the eigenvalue is actually the value of the constrianed maximum / minimum of f. Unless these two coincide (in which case $x^T Ax$ would have to be constant on the sphere, i.e., A would have to be a multiple of the unit matrix I), we indeed have found two different solutions λ (one for the min and one for the max).

(c) If A has an eigenvalue λ (with eigenvector $x \neq 0$), i.e., if $Ax = \lambda x$, then $x^T Ax = \lambda\|x\|^2$. For a positive definite matrix A, this expression must be positive for *all* vectors $x \neq 0$, in particular for the eigenvector x Hence $\lambda > 0$. Similarly we can argue for positive semidefinite.

Hwk #39:

Think of the task of finding the absolute maximum of $x^2 + \frac{1}{2}y^2 + y^4 - xy$ on the set S given by $(x-1)^2 + |y + y^3| \leq 5$. The purpose of this problem is NOT that you would actually do calculations to FIND the maximum (which would require numerical methods). Rather, in preparation for such a search. I want you to use the Hessian to conclude that the maximum exists and is on the BOUNDARY of the set S.

The message here is: While modest problems can already lead to prohibitively complicated calculations that may need numerical tools, simple analytic arguments may still be able significantly to reduce the amount of labor in a numerical search.

Solution: First we note that an absolute maximum exists, b/c we have a continuous expression on a bounded and closed set. We argue that any absolute maximum in this case cannt be in the interior because the Hessian is not negative semidefinite anywhere. (We do not attempt to solve the equations from vanishing of the gradient, even though we might of course have tried this, too.

The Hessian is $\begin{bmatrix} 2 & -1 \\ -1 & 1+12y^2 \end{bmatrix}$. It is clearly positive definite everywhere by the Hurwitz test. So if a critical point were to be found in the interior of S, it would not be an absolute maximum (but a relative minimum at least; possibly an absolute minimum).

Hwk #40:

You may or may not have seen the following formula (called Heron's formula): The area of a triangle with sides a, b, c is $\sqrt{s(s-a)(s-b)(s-c)}$ where s is the semiperimeter $\frac{1}{2}(a+b+c)$. Show that among all triangles with a given perimeter $2s = a + b + c$, the area takes an absolute minimum exactly for the equilateral triangle. (Explain first why an absolute minimum exists before calclulating it.)

Solution: Since $a, b, c \geq 0$ and $a + b + c = 2s$, clearly, the admissible choices of a, b, c lie in a bounded set. The set is also closed (because it is given by equations and non-strict inequalities for