Solutions to Homework 4

Rudin, page 239/15:

a) We have $(x^4+y^2)^2 - 4x^4y^2 = (x^4-y^2)^2 \geq 0$ for all $(x,y) \in \mathbf{R}^2$.

To see that f is continuous, it's enough to check continuity at $(0,0)$:

$$\lim_{(x,y)\to(0,0)} x^2+y^2-2x^2y - \frac{4x^6y^2}{(x^4+y^2)^2} = 0 - \lim_{(x,y)\to(0,0)} x^2 \frac{4x^4y^2}{(x^4+y^2)^2} = 0$$

since $x^2 \to 0$ whereas the magnitude of the other factor is bounded above by 1. Since $f(0,0)=0$ by definition, this shows that f is continuous at $(0,0)$.

b) A computation shows that

$$g_\theta(t) = t^2 - 2t^3(\cos^2\theta\sin\theta) - t^4h(t)$$

where

$$h(t) = \frac{4\cos^6\theta\sin^2\theta}{(t^2\cos^4\theta - \sin^2\theta)}$$

is defined and C^∞ even at $t=0$ (when $\theta \neq 2n\pi$, we have $\sin\theta \neq 0$ and this is clear; when $\theta = 2n\pi$ the numerator vanishes altogether and $h(t)\equiv 0$). Thus

$$\begin{aligned} g'_\theta(t) &= 2t - 6t^2(\cos^2\theta\sin\theta) - t^3(4h(t)+th'(t)) \\ g''_\theta(t) &= 2 - 12t(\cos^2\theta\sin\theta) - t^2(12h(t)+8th'(t)+t^2h''(t)) \end{aligned}$$

It follows that $g'_\theta(0)=0$, $g''_\theta(0)=2$. Thus $g_\theta(t)$ has a strict local minimum at $t=0$.

However, it doesn't follow that f has a local minimum at $(0,0)$. In fact,

$$f(x,x^2) = x^2+x^4-2x^4-\frac{4x^{10}}{(2x^4)^2} = x^2-x^4+x^2 = -x^4,$$

so that e.g. $\{(1/n,1/n^2)\}$ is a sequence of points converging to $(0,0)$ such that $f(1/n,1/n^2) < 0 = f(0,0)$ for every $n \in \mathbf{N}$.

Rudin, page 239/21b: We'll need the derivative:

$$f'(x,y) = (6x^2-6x, 6y^2+6y) = 6(x(x-1), y(y+1))$$

For functions $f:\mathbf{R}^2 \to \mathbf{R}$, the implicit function theorem allows us to solve $f(x,y)=0$ locally for x in terms of y provided the partial derivative of f with respect to x does not vanish—i.e. provided $x \neq 0,1$. Let us find out where on the zero level set of f this happens:

$$\begin{aligned} 0 &= f(0,y) = 2y^3+3y^2 \Rightarrow y = 0, -3/2. \\ 0 &= f(1,y) = -1+2y^3+3y^2 \Rightarrow y = -1, 1/2 \end{aligned}$$

So by the implicit function theorem, we can solve locally for x in terms of y except near the points $(0,0),(0,-3/2),(1,-1),(1,1/2)$.

Now if the goal is to solve for y in terms of x, then the implicit function theorem allows us to do so at any point on the zero level set where the partial derivative of f with respect to y vanishes. The problem points occur when $y=0$ or $y=-1$, and $2x^3-3x^2+2y^3+3y^2$. More specifically we can solve for y in terms of x except at $(0,3/2),(0,0),(1,-1),(-1/2,-1)$.

Rudin, page 239/23: The linear approximation of $f(x,y_1,y_2)$ about the point $(0,1,-1)$ is

$$L(x,y_1,y_2)=f(0,1,-1)+f'(0,1,-1)\begin{pmatrix} x \\ y_1-1 \\ y_2+1 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}+\begin{pmatrix} 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y_1-1 \\ y_2+1 \end{pmatrix}=x+y_2-1.$$

One way to restate the implicit function theorem is to say that you can locally solve $f=0$ for some of the variables in terms of the others if you can do the same for the linear approximation. So that's we do here—we want to solve $f(x,y_1,y_2)=0$ for x in terms of y_1 and y_2, so we replace f by L and get

$$x+y_2-1=0\Rightarrow x=1-y_2,$$

where the right side is the linear approximation of the implicit function $x=g(y_1,y_2)$. In particular,

$$\frac{\partial g}{\partial y_1}(1,-1)=0,\quad \frac{\partial g}{\partial y_2}(1,-1)=-1.$$

Supplementary problem 1: Since we'll need it in all parts of the problem, let us first note that

$$f'(x,y,z,w)=\begin{pmatrix} 2x & -1 & 1 & 0 \\ -1 & 2y & 0 & 1 \end{pmatrix}.$$

- In particular,

  $$f'(1,1,0,0)=\begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix}.$$

  Since the leftmost 2×2 submatrix of this matrix is invertible (its determinant is 3), the implicit function theorem guarantees us neighborhoods $U\ni(1,1)$, $V\ni(0,0)$ and a C^1 function $g:V\to U$ such that

  $$\{(x,y,z,w)\in V\times U : f(x,y,z,w)=(0,0)\}=\{(x,y,z,w)\in V\times U : (x,y)=g(z,w)\}.$$

  That is, we can solve $f(x,y,z,w)=(0,0)$ locally near $(1,1,0,0)$ for x and y in terms of z and w by setting $(x,y)=g(z,w)$.

- The obvious initial guess for a solution (x,y) of $f(x,y,.1,-.2)=(0,0)$ is $(x_0,y_0)=(1,1)$. So let us improve this guess by replacing f by its linear approximation about

$(1,1,.1,-.2)$ setting it equal to zero and solving for x and y while (z,w) is fixed at $(.1,-.2)$: i.e.

$$\begin{aligned}
\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= f(1,1,.1,-.2) + f'(1,1,.1,-.2)\begin{pmatrix} x_1-1 \\ y_1-1 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} .1 \\ -.2 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1-1 \\ y_1-1 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2x_1-y_1-.9 \\ 2y_1-x_1-1.2 \end{pmatrix}
\end{aligned}$$

which implies that $(x_1,y_1) = (1,1.1)$. Note that this actually does improve on our starting guess: we had $f(x_0,y_0,.1,-.2) = (.1,-.2)$ whereas $f(x_1,y_1,.1,-.2) = (0,.01)$.

- To get an even better guess (x_2,y_2), I repeat the above, using the linear approximation of f about $(x_1,y_1,.1,-.2)$.

$$\begin{aligned}
\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= f(1,1.1,.1,-.2) + f'(1,1.1,.1,-.2)\begin{pmatrix} x_2-1 \\ y_2-1.1 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ .01 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 2.2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_2-1 \\ y_2-1.1 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2x_2-y_2-.9 \\ 2.2y_2-x_2-1.41 \end{pmatrix}
\end{aligned}$$

which implies that $(x_2,y_2) = (.9970588\ldots,1.0941176\ldots)$. Plugging this guess into f gives

$$f(x_2,y_2,.1,-.2) \approx (8.7\times10^{-6}, 3.5\times10^{-5}).$$

Not bad, if I do say so.

Supplementary problem 2: The first conclusion of the implicit function theorem tells us that $f(x,y) = f(a,b)$ for (x,y) in a neighborhood W of (a,b) if and only if $x = g(y)$. In particular, $f(a,b) = f(a,b)$, so it must be that $a = g(b)$.

Given then that g is C^1, its linear approximation at b has the form

$$L_g(y) = g(b) + g'(b)(y-b) = a + g'(b)(y-b)$$

where we compute $g'(b)$ using the Chain Rule as follows. Let $H : V \to \mathbf{R}^{n+m}$ ($V \subset \mathbf{R}^m$ is the neighborhood of b specified in the theorem) be the function $H(y) = (g(y), y)$. Then we have

$$f \circ H(y) = f(g(y), y) = f(a, b)$$

for all $y \in V$. So by the Chain Rule,

$$f'(H(y)) \cdot H'(y) = 0$$

on V—in particular at $y = b$. Now recall that

$$f' = (D_x f, D_y f)$$

where $D_x f : \mathbf{R}^n \to \mathbf{R}^n$ and $D_y f : \mathbf{R}^m \to \mathbf{R}^n$ are the portions of f' corresponding to partial derivatives with respect to the x variables and y variables, respectively. Moreover,

$$H'(y) = \begin{pmatrix} g'(y) \\ \text{id} \end{pmatrix}.$$

where $g'(y) : \mathbf{R}^m \to \mathbf{R}^n$ and $\text{id} : \mathbf{R}^m \to \mathbf{R}^m$. Plugging these things back in gives

$$0 = (D_x f(H(b)), D_y f(H(b))) \cdot \begin{pmatrix} g'(b) \\ \text{id} \end{pmatrix} = (D_x f(a, b), D_y f(a, b)) \cdot \begin{pmatrix} g'(b) \\ \text{id} \end{pmatrix} = D_x f(a, b) g'(b) + D_y f(a, b).$$

Solving for $g'(b)$, we arrive at

$$g'(b) = -D_x f(a, b)^{-1} D_y f(a, b)$$

and

$$L_g(y) = a - D_x f(a, b)^{-1} D_y f(a, b)(y - b).$$

□