Problem 1

Let \( x_n = \varepsilon_n + p \). Then

\[
\varepsilon_{n+1} + p = x_{n+1} = g(x_n).
\]

If we taylor expand \( g(x_n) \) around \( p \), so that \( g(x_n) = g(p) + \varepsilon_n g'(p) + \mathcal{O}(\varepsilon_n^2) \), then we have

\[
\varepsilon_{n+1} + p = g(p) + \varepsilon_n g'(p) + \mathcal{O}(\varepsilon_n^2)
\]

but since \( g(p) = p \), we are left with

\[
\varepsilon_{n+1} = \varepsilon_n g'(p) + \mathcal{O}(\varepsilon_n^2).
\]

We can do a similar thing for \( x_{n+1} \):

\[
\varepsilon_{n+2} + p = x_{n+2} = g(x_{n+1}) = g(p) + \varepsilon_{n+1} g'(p) + \mathcal{O}(\varepsilon_{n+1}^2)
\]

\[
\Rightarrow \varepsilon_{n+2} = [\varepsilon_n g'(p) + \mathcal{O}(\varepsilon_n^2)] g'(p) + \mathcal{O}\left([\varepsilon_n g'(p) + \mathcal{O}(\varepsilon_n^2)]^2\right)
\]

\[
= \varepsilon_n [g'(p)]^2 + \mathcal{O}\left(\varepsilon_n^2 [g'(p)] + \varepsilon_n \mathcal{O}(\varepsilon_n^2) + [\mathcal{O}(\varepsilon_n^2)]^2\right)
\]

\[
= \varepsilon_n [g'(p)]^2 + \mathcal{O}(\varepsilon_n^2).
\]

For the next iterate, \( x_{n+2} \), we use Aitken extrapolation, so that

\[
\tilde{\varepsilon}_{n+2} + p = \tilde{x}_{n+2} = \text{Aitken}(x_{n+2})
\]

\[
= \frac{(x_{n+2} - x_{n+1})^2}{x_{n+2} - 2x_{n+1} + x_n}
\]

\[
= \varepsilon_{n+2} + p - \frac{[\varepsilon_{n+2} + p - (\varepsilon_{n+1} + p)]^2}{\varepsilon_{n+2} + p - 2(\varepsilon_{n+1} + p) + \varepsilon_n + p}
\]

\[
= \varepsilon_{n+2} + p - \frac{[\varepsilon_{n+2} - \varepsilon_n + p]^2}{\varepsilon_{n+2} - 2\varepsilon_{n+1} + \varepsilon_n}
\]

that is,

\[
\varepsilon_{n+2} = \varepsilon_{n+2} - \frac{[\varepsilon_{n+2} - \varepsilon_{n+1}]^2}{\varepsilon_{n+2} - 2\varepsilon_{n+1} + \varepsilon_n}
\]

Note that the first \( \varepsilon_{n+2} \) is just \( \varepsilon_n [g'(p)]^2 + \mathcal{O}(\varepsilon_n^2) \). To evaluate the second term, we can use Mathematica:
\[ \text{en2} = \text{en} (g'[p])^2 + O(\text{en})^2; \]
\[ \text{en1} = \text{en} g'[p] + O(\text{en})^2; \]

\[ \frac{\text{en2}^2 - 2 \text{en2 en1} + \text{en1}^2}{\text{en2} - 2 \text{en1} + \text{en}} \]

\[ \frac{(g'[p])^2 - 2 g'[p] + g'[p]^4}{1 - 2 g'[p] + g'[p]^2} \text{en} + O(\text{en})^2 \]

\[ \text{Simplify[\%]} \]
\[ -g'[p]^2 \text{en} + O(\text{en})^2 \]

Thus, we are left with
\[ \tilde{e}_{n+2} = e_n \left[ g'(p) \right]^2 + O(e_n^2) - [g'(p)]^2 + O(e_n^2) \]
\[ = O(e_n^2) \]

Therefore, since \( \tilde{e}_{n+2} = O(e_n^2) \), we see that Steffensen's Method is quadratically convergent.

\[ \square \]

Problem 2

Let's define
\[ g(x) = g \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \sqrt{1 + (x+y)^2} - \frac{2}{3} \\ \frac{1}{\sqrt{2}} \sqrt{1 + (x-y)^2} - \frac{2}{5} \end{pmatrix} = \begin{pmatrix} g_1(x) \\ g_2(x) \end{pmatrix}. \]

Then we see that the Jacobian \( G \) of \( g \) is
\[ G(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \frac{x+y}{\sqrt{1 + (x+y)^2}} & \frac{1}{\sqrt{2}} \frac{x+y}{\sqrt{1 + (x+y)^2}} \\ \frac{1}{\sqrt{2}} \frac{x-y}{\sqrt{1 + (x-y)^2}} & \frac{1}{\sqrt{2}} \frac{x-y}{\sqrt{1 + (x-y)^2}} \end{pmatrix} \]

Now, by Atkinson page 105, we want to find a convex domain \( D \) such that \( g(D) \subset D \), and
\[ \lambda = \max_{x \in D} \| G(x) \|_\infty < 1. \]

Let us start with the second requirement. Note that
\[ \lambda = \max_{x \in D} \| G(x) \|_\infty = \max \max_{x \in D} \sum_{i=1}^n |a_{i,k}| \]

\[ = \max \max_{x \in D} \left\{ 2 \left| \frac{x+y}{\sqrt{2} \sqrt{1 + (x+y)^2}} \right|, 2 \left| \frac{x-y}{\sqrt{2} \sqrt{1 + (x-y)^2}} \right| \right\} \]

\[ = \max \max_{x \in D} \left\{ \sqrt{2} \left| \frac{x+y}{\sqrt{1 + (x+y)^2}}, \sqrt{2} \left| \frac{x-y}{\sqrt{1 + (x-y)^2}} \right| \right\} \]

want \( \lambda < 1 \).
Splitting this into the two cases requires that both terms in the set must be less than 1.

Case 1: We see that

\[
\frac{\sqrt{2|x + y|}}{\sqrt{1 + (x + y)^2}} < 1 \implies |x + y| < \sqrt{\frac{1 + (x + y)^2}{2}}
\]

\[
\implies (x + y)^2 < \frac{1 + (x + y)^2}{2}
\]

\[
\implies \frac{(x + y)^2}{2} < \frac{1}{2}
\]

\[
\implies \sqrt{(x + y)^2} < \sqrt{1}
\]

\[
\implies |x + y| < 1.
\]

Case 2: Using a very similar argument, it can be shown that

\[
\frac{\sqrt{2|x - y|}}{\sqrt{1 + (x - y)^2}} < 1 \implies |x - y| < 1.
\]

Thus, these two requirements lead to the following four equations:

\[
y < 1 - x \quad y > -1 - x
\]

\[
y > x - 1 \quad y < 1 + x
\]

We can use Mathematica to visualize this region:

\[
\text{RegionPlot}[	ext{Abs}[x - y] < 1 \&\& \text{Abs}[x + y] < 1, \{x, -1, 1\}, \{y, -1, 1\}]
\]
Now, we need to show that \( g(\mathcal{D}) \subseteq \mathcal{D} \). Let \( \left( \frac{x}{y} \right) \in \mathcal{D} \). Then
\[
|g_1(x) + g_2(y)| = \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x + y)^2} - \frac{2}{3} + \frac{1}{\sqrt{2}} \sqrt{1 + (x - y)^2} - \frac{2}{3} \right|
\leq \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x + y)^2} - \frac{2}{3} \right| + \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x - y)^2} - \frac{2}{3} \right|
\]

However, note that
\[
0 < |x + y| < 1 \implies 0 < (x + y)^2 < 1
\implies 1 < 1 + (x + y)^2 < 2
\implies 1 < \sqrt{1 + (x + y)^2} < \sqrt{2}
\implies \frac{1}{\sqrt{2}} < \frac{\sqrt{1 + (x + y)^2}}{2} < 1
\implies \frac{1}{\sqrt{2}} - \frac{2}{3} < \frac{\sqrt{1 + (x + y)^2}}{2} - \frac{2}{3} < \frac{1}{3}
\implies \left| \frac{\sqrt{1 + (x + y)^2}}{2} - \frac{2}{3} \right| < \max \left\{ \frac{1}{\sqrt{2}} - \frac{2}{3}, \frac{1}{3} \right\} = \frac{1}{3}
\]

We can use the same reasoning to show that
\[
\frac{1}{\sqrt{2}} \sqrt{1 + (x - y)^2} - \frac{2}{3} < \frac{1}{3}
\]
yielding
\[
|g_1(x) + g_2(y)| < \frac{1}{3} + \frac{1}{3} < 1.
\]

Moreover,
\[
|g_1(x) - g_2(y)| = \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x + y)^2} - \frac{2}{3} + \frac{1}{\sqrt{2}} \sqrt{1 + (x - y)^2} - \frac{2}{3} \right|
\leq \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x + y)^2} - \frac{2}{3} \right| + \left| \frac{1}{\sqrt{2}} \sqrt{1 + (x - y)^2} - \frac{2}{3} \right|
\leq \frac{1}{3} + \frac{1}{3} \quad \text{(as has already been shown)}
\]
\[< 1.\]

Thus, since \( g(x) \in \{(x, y) : |x + y| < 1 \text{ and } |x - y| < 1\} = \mathcal{D} \), we have shown that \( g(\mathcal{D}) = \mathcal{D} \).
Thus, since \( \mathcal{D} \) is a convex domain, we have found a region for which the fixed-point iteration is guaranteed to converge. \( \square \)
Problem 3

(a) Recall that the gradient of $f(x, y)$ is

$$\nabla f(x, y) = \left( \frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right).$$

Thus, the gradient line through $(x_n, y_n)$ will be

$$y - y_n = \frac{f_y}{f_x}(x - x_n) \implies f_x(y - y_n) = f_y(x - x_n)$$

$$\implies f_x(y - y_n) - f_y(x - x_n) = 0.$$ 

Hence, define $f_1(x, y) = f_x(y - y_n) - f_y(x - x_n) = 0$, and $f_2(x, y) = f(x, y) = 0$. Now for Newton, we are looking for a scheme of the form

$$x_{n+1} = x_n - J^{-1}(x_n)f(x_n)$$

so we need to compute the Jacobian.

$$J(x_n) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} -f_y & f_x \\ f_x & f_y \end{pmatrix}$$

Then we see that

$$J^{-1} = \frac{1}{\det J}\begin{pmatrix} f_y & -f_x \\ -f_x & -f_y \end{pmatrix}$$

$$= \frac{1}{f_x^2 + f_y^2}\begin{pmatrix} f_y & -f_x \\ -f_x & -f_y \end{pmatrix}$$

$$= \frac{1}{f_x^2 + f_y^2}\begin{pmatrix} -f_y & f_x \\ f_x & f_y \end{pmatrix}.$$ 

Thus, our scheme is

$$x_{n+1} = x_n - \frac{1}{f_x^2 + f_y^2}\begin{pmatrix} -f_y & f_x \\ f_x & f_y \end{pmatrix}\begin{pmatrix} f \\ f \end{pmatrix} \begin{pmatrix} f \\ f \end{pmatrix}$$

Since each time we construct $f_1$ to be 0 by the proper choice of our next iterate, and $f(x_n, y_n)$ is not necessarily zero at a given iterate, so we just leave $f_2 = f$. But then

$$x_{n+1} = x_n - \frac{1}{f_x^2 + f_y^2}\begin{pmatrix} f & f_x \\ f & f_y \end{pmatrix}\begin{pmatrix} f \\ f \end{pmatrix}$$

$$= \begin{pmatrix} x - df_x \\ y - df_y \end{pmatrix}$$

which is exactly the scheme to be derived.

(b) It seems that the scheme should generalize to

$$\begin{cases} x_{n+1} = x_n - df_x \\ y_{n+1} = y_n - df_y \end{cases}$$

$$z_{n+1} = z_n - df_z$$

where now $d = \frac{1}{f_x^2 + f_y^2 + f_z^2}$. Indeed, we can implement this in Mathematica to see if we get evidence of quadratic convergence. The code is shown below:
Looking at the output, where column 1 corresponds to $x_n$, 2 to $y_n$, and 3 to $z_n$, we observe that

Since each new iterate has twice the correct digits as the previous one, we conclude that our scheme indeed converges quadratically. □

**Problem 4**

The following Mathematica code defines the functions $f_1, \ldots, f_6$ via the Sturm sequence technique:
Applying these functions at \( x = -2 \) and \( x = 3 \) yields

\[
\text{fs1} = \{f1[-2], f2[-2], f3[-2], f4[-2], f5[-2], f6[-2]\}
\]

Out[123]= \[24, -38, \frac{21}{8}, \frac{27200}{361}, -\frac{361}{400}, 0\]

\[
\text{fs2} = \{f1[3], f2[3], f3[3], f4[3], f5[3], f6[3]\}
\]

Out[124]= \[34, 67, 193, -\frac{12800}{361}, -\frac{361}{400}, 0\]

We see that there are 3 total sign changes for \( x = -2 \), and 1 total sign change for \( x = 3 \), so there must be \( |3 - 1| = 2 \) roots in \([-2,3]\). Applying these functions at \( x = -100 \) and \( x = 100 \) (noting that \( f \) is increasing without bound by the time it reaches \( \pm 100 \)) yields

\[
\text{fs1} = \{f1[-100], f2[-100], f3[-100], f4[-100], f5[-100], f6[-100]\}
\]

Out[125]= \[1009800204, -40292022, \frac{93369}{8}, \frac{8111200}{361}, -\frac{361}{400}, 0\]

\[
\text{fs2} = \{f1[100], f2[100], f3[100], f4[100], f5[100], f6[100]\}
\]

Out[126]= \[98979804, 3969598, \frac{96369}{8}, -\frac{798800}{361}, -\frac{361}{400}, 0\]

Here, there are 3 sign changes for \( x = -100 \) and 1 sign change for \( x = 100 \), so again there must be 2 roots on \((-\infty, \infty)\). Thus, these roots must be the same as those which occur in \([-2,3]\). We can verify this by plotting the function in Mathematica: