Minimal distance from a point to polynomial functions of degree 3 or less

Martin Thoma

January 2, 2014

Introduction

When you want to develop a selfdriving car, you have to plan which path it should take. A reasonable choice for the representation of paths are cubic splines. You also have to be able to calculate how to steer to get or to remain on a path. A way to do this is by applying the PID algorithm. This algorithm needs to know the signed current error. So you need to be able to get the minimal distance of a point (the position of the car) to a cubic spline (the prefered path) combined with sign (which represents the steering direction). As one steering direction might be prefered, it is not only necessary to get the minimal absolute distance, but might also help to get all points on the spline with minimal distance.

In this paper, I want to discuss how to find all points on a cubic function with minimal distance to a given point. As other representations of paths might be easier to understand and to implement, I will also cover the problem of finding the minimal distance of a point to a polynomial of degree 0, 1 and 2.

While I analyzed this problem, I've got interested in variations of the underlying PID-related problem. So I will try to give robust and easy-to-implement algorithms to calculate the distance of a point to a (piecewise or global) defined polynomial function of degree ≤ 3.

When you're able to calculate the distance to a polynomial which is defined on a closed invervall, you can calculate the distance from a point to a spline by calculating the distance to the pieces of the spline.

Contents

1 Description of the Problem

Let $f : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}$ be a polynomial function and $P \in \mathbb{R}^2$ be a point. Let $d_{P,f} : \mathbb{R} \to \mathbb{R}_0^+$ be the Euklidean distance of P to a point $(x, f(x))$ on the graph of f:

$$d_{P,f}(x) := \sqrt{(x - x_P)^2 + (f(x) - y_P)^2}$$

Now there is finite set $M = \{ x_1, \dots, x_n \} \subseteq D$ of minima for given f and P:

$$M = \left\{ x \in D \;\middle|\; d_{P,f}(x) = \min_{\overline{x} \in D} d_{P,f}(\overline{x}) \right\}$$

But minimizing $d_{P,f}$ is the same as minimizing $d_{P,f}^2 = (x_p^2 - 2x_p x + x^2) + (y_p^2 - 2y_p f(x) + f(x)^2)$.

In order to solve the minimal distance problem, Fermat's theorem about stationary points will be tremendously usefull:

Theorem 1 (Fermat's theorem about stationary points)
Let x_0 be a local extremum of a differentiable function $f : \mathbb{R} \to \mathbb{R}$.

Then: $f'(x_0) = 0$.

So in fact you can calculate the roots of $(d_{P,f}(x))'$ or $(d_{P,f}(x)^2)'$ to get candidates for minimal distance. $(d_{P,f}(x)^2)'$ is a polynomial if f is a polynomial. So if f is a polynomial, we can always get a finite number of candidates by finding roots of $(d_{P,f}(x)^2)'$. But this gets difficult when f has degree 3 or higher as explained in Theorem 6. Another problem one has to bear in mind is that these candidates include all points with minimal distance, but might also contain more. Example 1 shows such a situation.

Let S_n be the function that returns the set of solutions for a polynomial f of degree n and a point P:

$$S_n : \{ \text{Polynomials of degree } n \text{ defined on } \mathbb{R} \} \times \mathbb{R}^2 \to \mathcal{P}(\mathbb{R})$$

$$S_n(f, P) := \arg\min_{x \in \mathbb{R}} d_{P,f}(x) = M$$

If possible, I will explicitly give this function.

2 Constant functions

2.1 Defined on $\mathbb{R}$

Lemma 2
Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) := c$ with $c \in \mathbb{R}$ be a constant function.

Then $(x_P, f(x_P))$ is the only point on the graph of f with minimal distance to P.

The situation can be seen in Figure 2.1.

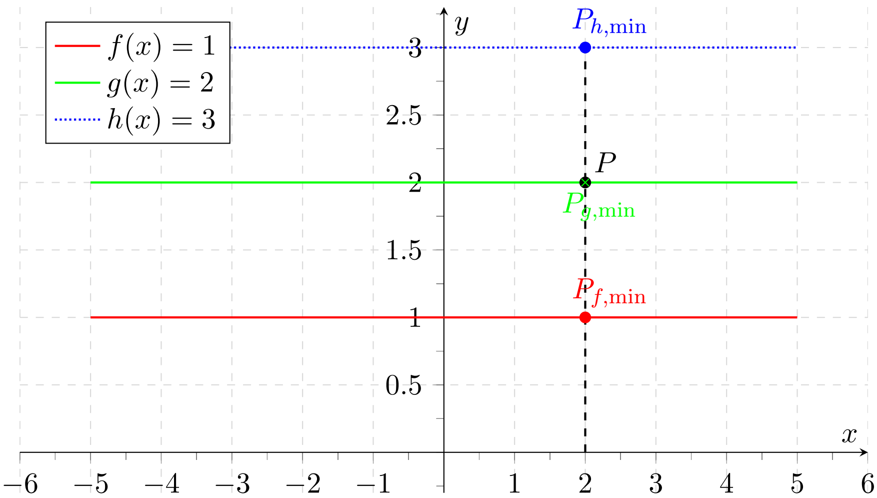

Figure 2.1: Three constant functions and their points with minimal distance

Proof: The point $(x, f(x))$ with minimal distance can be calculated directly:

$$
\begin{aligned}
d_{P,f}(x) &= \sqrt{(x - x_P)^2 + (f(x) - y_P)^2} && (2.1)\\
&= \sqrt{(x^2 - 2x_P x + x_P^2) + (c^2 - 2cy_P + y_P^2)} && (2.2)\\
&= \sqrt{x^2 - 2x_P x + (x_P^2 + c^2 - 2cy_P + y_P^2)} && (2.3)\\
\xRightarrow{\text{Theorem 1}} 0 &\stackrel{!}{=} (d_{P,f}(x)^2)' && (2.4)\\
&= 2x - 2x_P && (2.5)\\
\Leftrightarrow x &\stackrel{!}{=} x_P && (2.6)
\end{aligned}
$$

So $(x_P, f(x_P))$ is the only point with minimal distance to P. ■

This result means:

$$S_0(f, P) = \{\, x_P \,\} \text{ with } P = (x_P, y_P)$$

2.2 Defined on a closed interval $[a, b] \subseteq \mathbb{R}$

Theorem 3 (Solution formula for constant functions)
Let $f : [a, b] \to \mathbb{R}$, $f(x) := c$ with $a, b, c \in \mathbb{R}$ and $a \leq b$ be a constant function.

Then the point $(x, f(x))$ of f with minimal distance to P is given by:

$$\underset{x\in[a,b]}{\arg\min}\, d_{P,f}(x) = \begin{cases} S_0(f,P) & \text{if } S_0(f,P) \cap [a,b] \neq \emptyset \\ \{\, a \,\} & \text{if } S_0(f,P) \ni x_P < a \\ \{\, b \,\} & \text{if } S_0(f,P) \ni x_P > b \end{cases}$$

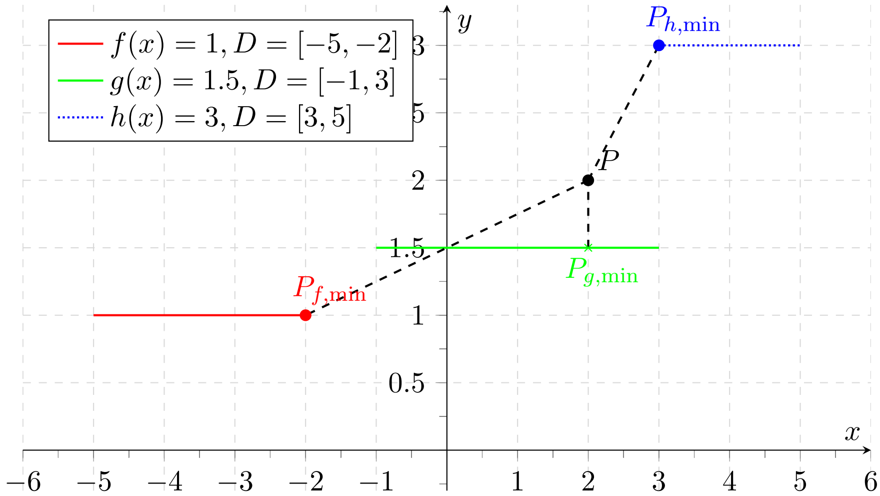

Figure 2.2: Three constant functions and their points with minimal distance

Proof:

$$\underset{x\in[a,b]}{\arg\min}\, d_{P,f}(x) = \underset{x\in[a,b]}{\arg\min}\, d_{P,f}(x)^2 \tag{2.7}$$

$$= \underset{x\in[a,b]}{\arg\min}\Big((x - x_P)^2 + \overbrace{(y_P^2 - 2y_P c + c^2)}^{\text{constant}}\Big) \tag{2.8}$$

$$= \underset{x\in[a,b]}{\arg\min}(x - x_P)^2 \tag{2.9}$$

which is optimal for $x = x_P$, but if $x_P \notin [a, b]$, you want to make this term as small as possible. It gets as small as possible when x is as similar to x_p as possible. This yields directly to the solution formula. ■

3 Linear function

3.1 Defined on $\mathbb{R}$

Theorem 4 (Solution formula for linear functions on $\mathbb{R}$)
Let $f : \mathbb{R} \to \mathbb{R}$ be a linear function $f(x) := m \cdot x + t$ with $m \in \mathbb{R} \setminus \{ 0 \}$ and $t \in \mathbb{R}$ be a linear function.

Then there is only one point $(x, f(x))$ on the graph of f with minimal distance to $P = (x_P, y_P)$. This point is given by

$$x = \frac{m}{m^2+1}\left(y_P + \frac{1}{m} \cdot x_P - t\right)$$

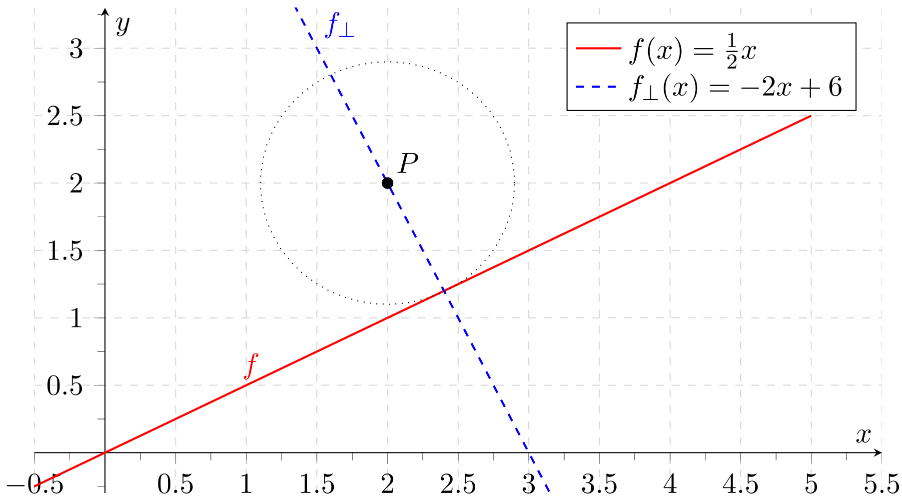

Figure 3.1: The shortest distance of P to f can be calculated by using the perpendicular

Proof: With Theorem 1 you get:

$$0 \stackrel{!}{=} (d_{P,f}(x)^2)' \tag{3.1}$$
$$= 2(x - x_P) + 2(f(x) - y_P)f'(x) \tag{3.2}$$
$$\Leftrightarrow 0 \stackrel{!}{=} x - x_P + (f(x) - y_P)f'(x) \tag{3.3}$$
$$= x - x_P + (mx + t - y_P) \cdot m \tag{3.4}$$
$$= x(m+1) + m(t - y_P) - x_P \tag{3.5}$$
$$\Leftrightarrow x \stackrel{!}{=} \frac{x_p - m(t - y_p)}{m^2+1} \tag{3.6}$$
$$= \frac{m}{m^2+1}\left(y_P + \frac{1}{m} \cdot x_P - t\right) \tag{3.7}$$

It is obvious that a minium has to exist, the x from Equation 3.7 has to be this minimum. ■

3.2 Defined on a closed interval $[a, b] \subseteq \mathbb{R}$

Let $f : [a, b] \to \mathbb{R}$, $f(x) := m \cdot x + t$ with $a, b, m, t \in \mathbb{R}$ and $a \leq b$, $m \neq 0$ be a linear function.

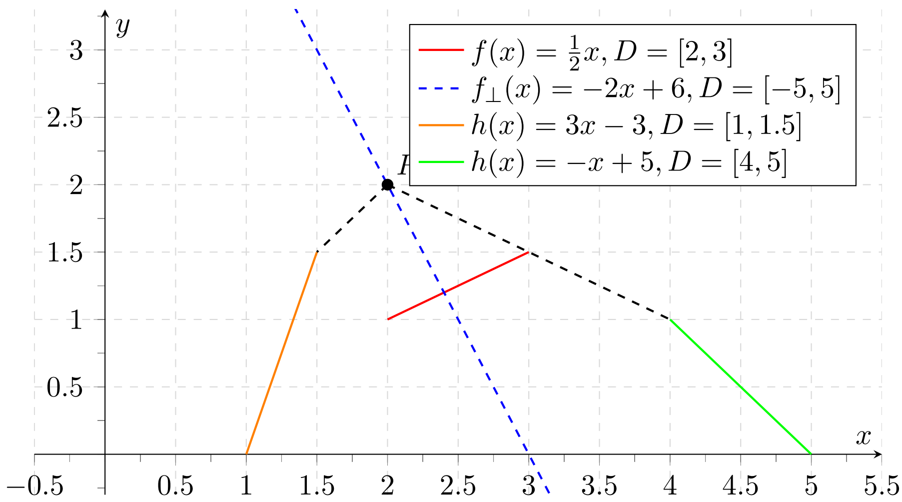

Figure 3.2: Different situations when you have linear functions which are defined on a closed intervall

The point with minimum distance can be found by:

$$\arg\min_{x \in [a,b]} d_{P,f}(x) = \begin{cases} S_1(f, P) & \text{if } S_1(f, P) \cap [a, b] \neq \emptyset \\ \{\, a \,\} & \text{if } S_1(f, P) \ni x < a \\ \{\, b \,\} & \text{if } S_1(f, P) \ni x > b \end{cases}$$

If $S_1(f, P) \cap [a, b] \neq \emptyset$, then $\arg\min_{x \in [a,b]} d_{P,f}(x) = S_1(f, P) \cap [a, b]$, because $S_1(f, P)$ gives all global minima of f. Those are also minima for the intervall $[a, b]$. There are not more minima, because S_1 gives all minima of P to f.

If $S_1(f, P) \cap [a, b] = \emptyset$, then it is not that simple. But we can calculate the distance function:

$$\begin{aligned} d_{P,f}(x) &= \sqrt{(x - x_P)^2 + (f(x) - y_P)^2} && (3.8) \\ &= \sqrt{(x^2 - 2xx_P + x_P^2) + (mx + (t - y_P))^2} && (3.9) \\ &= \sqrt{(x^2 - 2xx_P + x_P^2) + m^2x^2 + 2mx(t - y_P) + (t - y_P)^2} && (3.10) \\ &= \sqrt{x^2(1 + m^2) + x(-2x_P + 2m(t - y_P)) + (x_P^2 + (t - y_P)^2)} && (3.11) \end{aligned}$$

This function (defined on $\mathbb{R}$) is symmetry to the axis

$$\begin{aligned} x_S &= -\frac{-2x_P + 2m(t - y_P)}{2(1 + m^2)} && (3.12) \\ &= \frac{x_P - m(t - y_P)}{1 + m^2} && (3.13) \\ &= \frac{m}{m^2 + 1}(y_P + \frac{1}{m}x_P - t) && (3.14) \end{aligned}$$

f is on $(-\infty, x_S]$ strictly monotonically decreasing and on $[x_S, +\infty)$ strictly monotonically increasing.

Thus we can conclude:

$$\forall x, y \in \mathbb{R} : x \leq y < x_S \Rightarrow d_{P,f}(x_S) < d_{P,f}(y) \leq d_{P,f}(x)$$

$$\forall x, y \in \mathbb{R} : x_S < y \leq x \Rightarrow d_{P,f}(x_S) < d_{P,f}(y) \leq d_{P,f}(x)$$

When $S_1(f, P) \cap [a, b] = \emptyset$, then you can have two cases:

- $a \leq b < x_S$: b has the shortest distance in $[a, b]$ on the graph of f to P.
- $x_S < a \leq b$: a has the shortest distance in $[a, b]$ on the graph of f to P.

4 Quadratic functions

4.1 Defined on $\mathbb{R}$

Let $f : \mathbb{R} \to \mathbb{R}, f(x) = a \cdot x^2 + b \cdot x + c$ with $a \in \mathbb{R} \setminus \{ 0 \}$ and $b, c \in \mathbb{R}$ be a quadratic function.

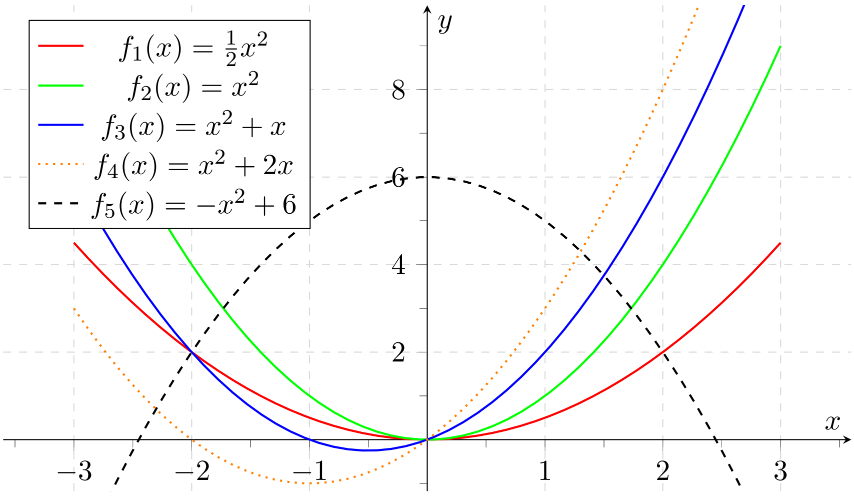

Figure 4.1: Quadratic functions

4.1.1 Calculate points with minimal distance

In this case, $d_{P,f}^2$ is polynomial of degree $n^2 = 4$. We use Theorem 1:

$$\begin{aligned}
0 &\stackrel{!}{=} (d_{P,f}^2)' && (4.1)\\
&= 2x - 2x_p - 2y_p f'(x) + \left(f(x)^2\right)' && (4.2)\\
&= 2x - 2x_p - 2y_p f'(x) + 2f(x) \cdot f'(x) \qquad \text{(chain rule)} && (4.3)\\
\Leftrightarrow 0 &\stackrel{!}{=} x - x_p - y_p f'(x) + f(x) \cdot f'(x) \qquad \text{(divide by 2)} && (4.4)\\
&= x - x_p - y_p(2ax + b) + (ax^2 + bx + c)(2ax + b) && (4.5)\\
&= x - x_p - y_p \cdot 2ax - y_p b + (2a^2x^3 + 2abx^2 + 2acx + abx^2 + b^2x + bc) && (4.6)\\
&= x - x_p - 2y_p ax - y_p b + (2a^2x^3 + 3abx^2 + 2acx + b^2x + bc) && (4.7)\\
&= 2a^2x^3 + 3abx^2 + (1 - 2y_p a + 2ac + b^2)x + (bc - by_p - x_p) && (4.8)
\end{aligned}$$

This is an algebraic equation of degree 3. There can be up to 3 solutions in such an equation. Those solutions can be found with a closed formula. But not every solution of the equation given by Theorem 1 has to be a solution to the given problem as you can see in Example 1.

Example 1

Let $a = 1, b = 0, c = -1, x_p = 0, y_p = 1$. So $f(x) = x^2 - 1$ and $P(0, 1)$.

$$\begin{aligned}
\xrightarrow{\text{Equation 4.8}} 0 &\overset{!}{=} 2x^3 - 3x && (4.9)\\
&= x(2x^2 - 3) && (4.10)\\
\Rightarrow x_{1,2} &= \pm\sqrt{\frac{3}{2}} \text{ and } x_3 = 0 && (4.11)\\
d_{P,f}(x_3) &= \sqrt{0^2 + (-1-1)^2} = 2 && (4.12)\\
d_{P,f}\left(\pm\sqrt{\frac{3}{2}}\right) &= \sqrt{\left(\sqrt{\frac{3}{2}} - 0\right)^2 + \left(\frac{1}{2} - 1\right)^2} && (4.13)\\
&= \sqrt{3/2 + 1/4} && (4.14)\\
&= \sqrt{7/4} && (4.15)\\
& && (4.16)
\end{aligned}$$

This means x_3 is not a point of minimal distance, although $(d_{P,f}(x_3))' = 0$.

4.1.2 Number of points with minimal distance

Theorem 5

A point P has either one or two points on the graph of a quadratic function f that are closest to P.

Proof: The number of closests points of f cannot be bigger than 3, because Equation 4.8 is a polynomial function of degree 3. Such a function can have at most 3 roots. As f has at least one point on its graph, there is at least one point with minimal distance.

In the following, I will do some transformations with $f = f_0$ and $P = P_0$. This will make it easier to calculate the minimal distance points. Moving f_0 and P_0 simultaneously in x or y direction does not change the minimum distance. Furthermore, we can find the points with minimum distance on the moved situation and calculate the minimum points in the original situation.

First of all, we move f_0 and P_0 by $\frac{b}{2a}$ in x direction, so

$$f_1(x) = ax^2 - \frac{b^2}{4a} + c \quad \text{and} \quad P_1 = \left(x_p + \frac{b}{2a},\ y_p\right)$$

Because:[1]

$$\begin{aligned}
f(x - b/2a) &= a(x - b/2a)^2 + b(x - b/2a) + c && (4.17)\\
&= a(x^2 - b/ax + b^2/4a^2) + bx - b^2/2a + c && (4.18)\\
&= ax^2 - bx + b^2/4a + bx - b^2/2a + c && (4.19)\\
&= ax^2 - b^2/4a + c && (4.20)
\end{aligned}$$

[1] The idea why you subtract $\frac{b}{2a}$ within f is that when you subtract something from x before applying f it takes more time (x needs to be bigger) to get to the same situation. In consequence, if we want to move the whole graph by 1 to the left, we have to add +1.

Then move f_1 and P_1 by $\frac{b^2}{4a} - c$ in y direction. You get:

$$f_2(x) = ax^2 \quad \text{and} \quad P_2 = \Big(\underbrace{x_P + \frac{b}{2a}}_{=:z},\ \underbrace{y_P + \frac{b^2}{4a} - c}_{=:w}\Big)$$

As $f_2(x) = ax^2$ is symmetric to the y axis, only points $P = (0, w)$ could possilby have three minima.

Then compute:

$$\begin{aligned}
d_{P,f_2}(x) &= \sqrt{(x-0)^2 + (f_2(x) - w)^2} && (4.21)\\
&= \sqrt{x^2 + (ax^2 - w)^2} && (4.22)\\
&= \sqrt{x^2 + a^2x^4 - 2awx^2 + w^2} && (4.23)\\
&= \sqrt{a^2x^4 + (1-2aw)x^2 + w^2} && (4.24)\\
&= \sqrt{\left(ax^2 + \frac{1-2aw}{2a}\right)^2 + w^2 - \left(\frac{1-2aw}{2a}\right)^2} && (4.25)\\
&= \sqrt{(ax^2 + 1/2a - w)^2 + \left(w^2 - \left(\frac{1-2aw}{2a}\right)^2\right)} && (4.26)
\end{aligned}$$

This means, the term

$$a^2x^2 + (1/2a - w)$$

has to get as close to 0 as possilbe when we want to minimize d_{P,f_2}. For $w \leq 1/2a$ you only have $x = 0$ as a minimum. For all other points $P = (0, w)$, there are exactly two minima $x_{1,2} = \pm\sqrt{\frac{\frac{1}{2a} - w}{a}}$. ■

4.1.3 Solution formula

We start with the graph that was moved so that $f_2 = ax^2$.

Case 1: P is on the symmetry axis, hence $x_P = -\frac{b}{2a}$.

In this case, we have already found the solution. If $w = y_P + \frac{b^2}{4a} - c > \frac{1}{2a}$, then there are two solutions:

$$x_{1,2} = \pm\sqrt{\frac{\frac{1}{2a} - w}{a}}$$

Otherwise, there is only one solution $x_1 = 0$.

Case 2: $P = (z, w)$ is not on the symmetry axis, so $z \neq 0$. Then you compute:

$$\begin{aligned}
d_{P,f_2}(x) &= \sqrt{(x-z)^2 + (f(x)-w)^2} && (4.27)\\
&= \sqrt{(x^2 - 2zx + z^2) + ((ax^2)^2 - 2awx^2 + w^2)} && (4.28)\\
&= \sqrt{a^2x^4 + (1-2aw)x^2 + (-2z)x + z^2 + w^2} && (4.29)\\
0 &\stackrel{!}{=} \left(\left(d_{P,f_2}(x)\right)^2\right)' && (4.30)\\
&= 4a^2x^3 + 2(1-2aw)x + (-2z) && (4.31)\\
&= 2\left(2a^2x^3 + (1-2aw)x\right) - 2z && (4.32)\\
\Leftrightarrow 0 &\stackrel{!}{=} 2a^2x^3 + (1-2aw)x - z && (4.33)\\
\stackrel{a\neq 0}{\Leftrightarrow} 0 &\stackrel{!}{=} x^3 + \underbrace{\frac{1-2aw}{2a^2}}_{=:\alpha} x + \underbrace{\frac{-z}{2a^2}}_{=:\beta} && (4.34)\\
&= x^3 + \alpha x + \beta && (4.35)
\end{aligned}$$

Let t be defined as

$$t := \sqrt[3]{\sqrt{3 \cdot (4\alpha^3 + 27\beta^2)} - 9\beta}$$

I will make use of the following identities:

$$\begin{aligned}
(1 - i\sqrt{3})^2 &= -2(1 + i\sqrt{3})\\
(1 + i\sqrt{3})^2 &= -2(1 - i\sqrt{3})\\
(1 \pm i\sqrt{3})^3 &= -8\\
(a-b)^3 &= a^3 - 3a^2b + 3ab^2 - b^3
\end{aligned}$$

Case 2.1: $4\alpha^3 + 27\beta^2 \geq 0$: The first solution of $x^3 + \alpha x + \beta = 0$ is

$$x = \frac{t}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}\alpha}{t}$$

Let's validate this solution:

$$\begin{aligned}
0 &\stackrel{!}{=} \left(\frac{t}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}\alpha}{t}\right)^3 + \alpha\left(\frac{t}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}\alpha}{t}\right) + \beta && (4.36)\\
&= \left(\frac{t}{\sqrt[3]{18}}\right)^3 - 3\left(\frac{t}{\sqrt[3]{18}}\right)^2 \frac{\sqrt[3]{\frac{2}{3}}\alpha}{t} + 3\left(\frac{t}{\sqrt[3]{18}}\right)\left(\frac{\sqrt[3]{\frac{2}{3}}\alpha}{t}\right)^2 - \left(\frac{\sqrt[3]{\frac{2}{3}}\alpha}{t}\right)^3 + \frac{t\alpha}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}\alpha^2}{t} + \beta && (4.37)\\
&= \frac{t^3}{18} - \frac{3t^2}{\sqrt[3]{18^2}}\frac{\sqrt[3]{\frac{2}{3}}\alpha}{t} + \frac{3t}{\sqrt[3]{18}}\frac{\sqrt[3]{\frac{4}{9}}\alpha^2}{t^2} - \frac{\frac{2}{3}\alpha^3}{t^3} + \frac{t\alpha}{\sqrt[3]{18}} - \frac{\sqrt[3]{2}\alpha^2}{\sqrt[3]{3}t} + \beta && (4.38)\\
&= \frac{t^3}{18} - \frac{\sqrt[3]{18}t\alpha}{\sqrt[3]{18^2}} + \frac{\sqrt[3]{12}\alpha^2}{\sqrt[3]{18}t} - \frac{2\alpha^3}{3t^3} + \frac{t\alpha}{\sqrt[3]{18}} - \frac{\sqrt[3]{2}\alpha^2}{\sqrt[3]{3}t} + \beta && (4.39)\\
&= \frac{t^3}{18} - \frac{t\alpha}{\sqrt[3]{18}} + \frac{\sqrt[3]{2}\alpha^2}{\sqrt[3]{3}t} - \frac{2\alpha^3}{3t^3} + \frac{t\alpha}{\sqrt[3]{18}} - \frac{\sqrt[3]{2}\alpha^2}{\sqrt[3]{3}t} + \beta && (4.40)\\
&= \frac{t^3}{18} - \frac{2\alpha^3}{3t^3} + \beta && (4.41)
\end{aligned}$$

$$= \frac{t^6 - 12\alpha^3 + \beta 18 t^3}{18t^3} \tag{4.42}$$

Now only go on calculating with the numerator. Start with resubstituting t:

$$\begin{aligned}
0 &= (\sqrt{3 \cdot (4\alpha^3 + 27\beta^2)} - 9\beta)^2 - 12\alpha^3 + \beta 18(\sqrt{3 \cdot (4\alpha^3 + 27\beta^2)} - 9\beta) && (4.43)\\
&= (\sqrt{3 \cdot (4\alpha^3 + 27\beta^2)})^2 + (9\beta)^2 - 12\alpha^3 - (2 \cdot 9) \cdot 9\beta^2 && (4.44)\\
&= 3 \cdot (4\alpha^3 + 27\beta^2) - 81\beta^2 - 12\alpha^3 && (4.45)\\
&= 0 && (4.46)
\end{aligned}$$

Case 2.2: The second solution of $x^3 + \alpha x + \beta = 0$ is

$$x = \frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t} - \frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}$$

We will verify it in multiple steps. First, calculate x^3:

$$x^3 = \underbrace{\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right)^3}_{=:①} \underbrace{-3\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right)^2 \left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right)}_{=:②} \tag{4.47}$$

$$\underbrace{+3\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right) \left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right)^2}_{=:③} \underbrace{-\left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right)^3}_{=:④} \tag{4.48}$$

Now simplify the summands of x^3:

$$\begin{aligned}
① &= \left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right)^3 && (4.49)\\
&= \frac{-8\alpha^3}{12t^3} && (4.50)\\
&= \frac{-2\alpha^3}{3t^3} && (4.51)\\
② &= -3\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right)^2 \left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right) && (4.52)\\
&= \frac{-3\alpha^2(-2(1 - i\sqrt{3}))(1 - i\sqrt{3})t}{t^2\sqrt[3]{2^4 \cdot 3^2} \cdot 2\sqrt[3]{2 \cdot 3^2}} && (4.53)\\
&= \frac{6\alpha^2 t(-2(1 + i\sqrt{3}))}{12t^2\sqrt[3]{12}} && (4.54)\\
&= \frac{-\alpha^2(1 + i\sqrt{3})}{t\sqrt[3]{12}} && (4.55)\\
③ &= 3\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t}\right) \left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right)^2 && (4.56)\\
&= \frac{3\alpha t(1 + i\sqrt{3})(-2(1 + i\sqrt{3}))}{4 \cdot \sqrt[3]{12 \cdot 18^2}} && (4.57)\\
&= \frac{-\alpha t(-2(1 - i\sqrt{3}))}{2\sqrt[3]{12 \cdot 4 \cdot 3}} && (4.58)
\end{aligned}$$

$$= \frac{\alpha t(1 - i\sqrt{3})}{\sqrt[3]{2^4 \cdot 3^2}} \tag{4.59}$$

$$= \frac{\alpha t(1 - i\sqrt{3}}{2\sqrt[3]{18}} \tag{4.60}$$

$$④ = -\left(\frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right)^3 \tag{4.61}$$

$$= -\frac{(-8)t^3}{8 \cdot 18} \tag{4.62}$$

$$= \frac{t^3}{18} \tag{4.63}$$

Now get back to the original equation:

$$0 \overset{!}{=} x^3 + \alpha x + \beta \tag{4.64}$$

$$= \left(\frac{-2\alpha^3}{3t^3} + \frac{-\alpha^2(1 + \sqrt{3}i)}{t\sqrt[3]{12}} + \frac{\alpha t(1 - \sqrt{3}i)}{2\sqrt[3]{18}} + \frac{t^3}{18}\right) \tag{4.65}$$

$$+ \alpha\left(\frac{(1 + i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t} - \frac{(1 - i\sqrt{3})t}{2\sqrt[3]{18}}\right) + \beta \tag{4.66}$$

$$= \frac{-2\alpha^3}{3t^3} + \frac{t^3}{18} + \beta \tag{4.67}$$

$$= \frac{-12\alpha^3 + t^6 + 18t^3\beta}{18t^3} \tag{4.68}$$

Now continue with only the numerator

$$0 \overset{!}{=} -12\alpha^3 + (\sqrt{3(4\alpha^3 + 27\beta^2)} - 9\beta)^2 + 18(\sqrt{3(4\alpha^3 + 27\beta^2)} - 9\beta)\beta \tag{4.69}$$

$$= -12\alpha^3 + \left(3(4\alpha^3 + 27\beta^2) - 2 \cdot \sqrt{3(4\alpha^3 + 27\beta^2)} \cdot 9\beta + 81\beta^2\right) \tag{4.70}$$

$$+ 18\beta(\sqrt{3(4\alpha^3 + 27\beta^2)} - 9\beta) \tag{4.71}$$

$$= 81\beta^2 + 81\beta^2 - 2 \cdot 81\beta^2 \tag{4.72}$$

$$= 0 \tag{4.73}$$

Case 2.3: The third and thus last solution of $x^3 + \alpha x + \beta = 0$ is

$$x = \frac{(1 - i\sqrt{3})\alpha}{\sqrt[3]{12} \cdot t} - \frac{(1 + i\sqrt{3})t}{2\sqrt[3]{18}}$$

The complex conjugate root theorem states that if x is a complex root of a polynomial P, then its complex conjugate $\overline{x}$ is also a root of P. The solution presented in this case is the complex conjugate of case 2.2.

So the solution is given by

NO! Currently, there are erros in the solution. Check $f(x) = x^2$ and $P = (-2, 4)$. Solution should be $x_1 = -2$, but it isn't!

$$x_S := -\frac{b}{2a} \quad \text{(the symmetry axis)}$$

$$
\begin{aligned}
w &:= y_P + \frac{b^2}{4a} - c \quad \text{and} \quad z := x_P + \frac{b}{2a} \\
\alpha &:= \frac{1 - 2aw}{2a^2} \quad \text{and} \quad \beta := \frac{-z}{2a^2} \\
t &:= \sqrt[3]{\sqrt{3 \cdot (4\alpha^3 + 27\beta^2)} - 9\beta}
\end{aligned}
$$

$$
\arg\min_{x \in \mathbb{R}} d_{P,f}(x) = \begin{cases}
x_1 = +\sqrt{a(y_p + \frac{b^2}{4a} - c) - \frac{1}{2}} + x_S \text{ and} & \text{if } x_P = x_S \text{ and } y_p + \frac{b^2}{4a} - c > \frac{1}{2a} \\
x_2 = -\sqrt{a(y_p + \frac{b^2}{4a} - c) - \frac{1}{2}} + x_S & \\
x_1 = x_S & \text{if } x_P = x_S \text{ and } y_p + \frac{b^2}{4a} - c \leq \frac{1}{2a} \\
x_1 = \frac{t}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}\alpha}{t} & \text{if } x_P \neq x_S
\end{cases}
$$

4.2 Defined on a closed interval $[a, b] \subseteq \mathbb{R}$

Now the problem isn't as simple as with constant and linear functions.

If one of the minima in $S_2(P, f)$ is in $[a, b]$, this will be the shortest distance as there are no shorter distances.

The following IS WRONG! Can I include it to help the reader understand the problem?

If the function (defined on $\mathbb{R}$) has only one shortest distance point x for the given P, it's also easy: The point in $[a, b]$ that is closest to x will have the sortest distance.

$$\arg\min_{x \in [a,b]} d_{P,f}(x) = \begin{cases} S_2(f, P) \cap [a, b] & \text{if } S_2(f, P) \cap [a, b] \neq \emptyset \\ \{ a \} & \text{if } |S_2(f, P)| = 1 \text{ and } S_2(f, P) \ni x < a \\ \{ b \} & \text{if } |S_2(f, P)| = 1 \text{ and } S_2(f, P) \ni x > b \\ \textit{todo} & \text{if } |S_2(f, P)| = 2 \text{ and } S_2(f, P) \cap [a, b] = \emptyset \end{cases}$$

5 Cubic functions

5.1 Defined on $\mathbb{R}$

Let $f : \mathbb{R} \to \mathbb{R}, f(x) = a \cdot x^3 + b \cdot x^2 + c \cdot x + d$ be a cubic function with $a \in \mathbb{R} \setminus \{ 0 \}$ and $b, c, d \in \mathbb{R}$.

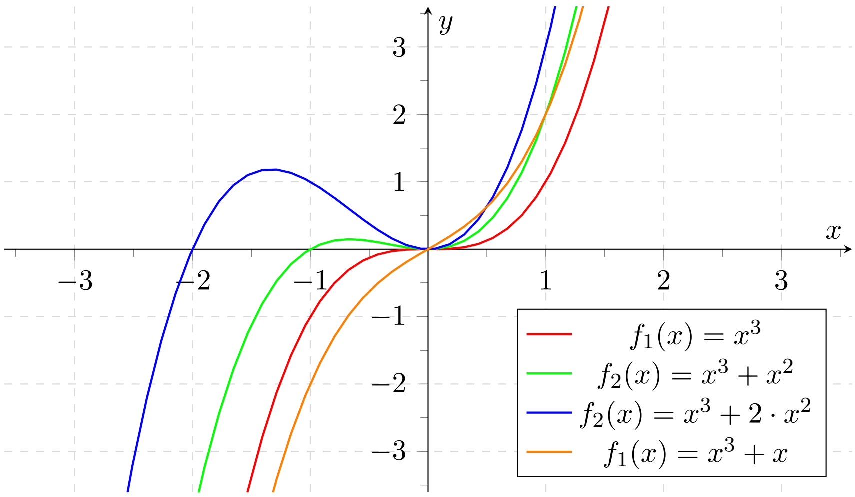

Figure 5.1: Cubic functions

5.1.1 Calculate points with minimal distance

Theorem 6
There cannot be a finite, closed form solution to the problem of finding a closest point $(x, f(x))$ to a given point P when f is a polynomial function of degree 3 or higher.

Proof: Suppose you could solve the closest point problem for arbitrary cubic functions $f = ax^3 + bx^2 + cx + d$ and arbitrary points $P = (x_P, y_P)$.

Then you could solve the following problem for x:

$$\begin{aligned}
0 &\overset{!}{=} \left((d_{P,f}(x))^2\right)' && (5.1)\\
&= -2x_p + 2x - 2y_p(f(x))' + (f(x)^2)' && (5.2)\\
&= 2f(x) \cdot f'(x) - 2y_p f'(x) + 2x - 2x_p && (5.3)\\
&= f(x) \cdot f'(x) - y_p f'(x) + x - x_p && (5.4)\\
&= \underbrace{f'(x) \cdot (f(x) - y_p)}_{\text{Polynomial of degree 5}} + x - x_p && (5.5)
\end{aligned}$$

General algebraic equations of degree 5 don't have a solution formula.[1] Although here seems

[1] TODO: Quelle

to be more structure, the resulting algebraic equation can be almost any polynomial of degree 5:[2]

$$0 \stackrel{!}{=} f'(x) \cdot (f(x) - y_p) + (x - x_p) \tag{5.6}$$
$$= \underbrace{3a^2}_{=\tilde{a}} x^5 + \underbrace{5ab}_{=\tilde{b}} x^4 + \underbrace{2(2ac + b^2)}_{=\tilde{c}} x^3 \quad + \underbrace{3(ad + bc - ay_p)}_{=\tilde{d}} x^2 \tag{5.7}$$
$$+ \underbrace{(2bd + c^2 + 1 - 2by_p)}_{=\tilde{e}} x + \underbrace{cd - cy_p - x_p}_{=\tilde{f}} \tag{5.8}$$
$$0 \stackrel{!}{=} \tilde{a}x^5 + \tilde{b}x^4 + \tilde{c}x^3 + \tilde{d}x^2 + \tilde{e}x + \tilde{f} \tag{5.9}$$

1. For any coefficient $\tilde{a} \in \mathbb{R}_{>0}$ of x^5 we can choose $a := \frac{1}{3}\sqrt{\tilde{a}}$ such that we get $\tilde{a}$.
2. For any coefficient $\tilde{b} \in \mathbb{R} \setminus \{ 0 \}$ of x^4 we can choose $b := \frac{1}{5a} \cdot \tilde{b}$ such that we get $\tilde{b}$.
3. With $c := -2b^2 + \frac{1}{4a}\tilde{c}$, we can get any value of $\tilde{c} \in \mathbb{R}$.
4. With $d := -bc + ay_p + \frac{1}{a}\tilde{d}$, we can get any value of $\tilde{d} \in \mathbb{R}$.
5. With $y_p := \frac{1}{2b}(2bd + c^2) \cdot \tilde{e}$, we can get any value of $\tilde{e} \in \mathbb{R}$.
6. With $x_p := cd - cy_P + \tilde{f}$, we can get any value of $\tilde{f} \in \mathbb{R}$.

The first restriction guarantees that we have a polynomial of degree 5. The second one is necessary, to get a high range of $\tilde{e}$.

This means that there is no finite solution formula for the problem of finding the closest points on a cubic function to a given point, because if there was one, you could use this formula for finding roots of polynomials of degree 5. ■

5.1.2 Another approach

Just like we moved the function f and the point to get in a nicer situation, we can apply this approach for cubic functions.

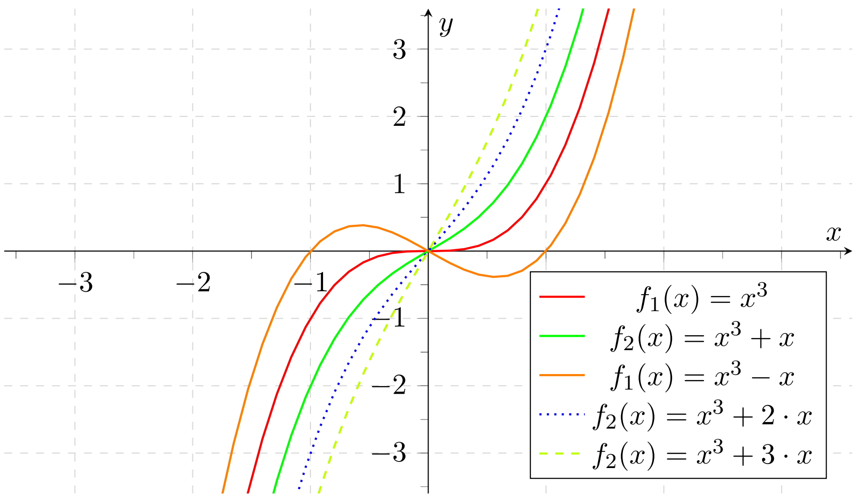

Figure 5.2: Cubic functions with $b = d = 0$

[2]Thanks to Peter Košinár on math.stackexchange.com for the idea.

First, we move f_0 by $\frac{b}{3a}$ in x direction, so

$$f_1(x) = ax^3 + \frac{b^2(c-1)}{3a}x + \frac{2b^3}{27a^2} - \frac{bc}{3a} + d \quad \text{and} \quad P_1 = (x_P + \frac{b}{3a}, y_P)$$

because

$$\begin{aligned}
f_1(x) &= a\left(x - \frac{b}{3a}\right)^3 + b\left(x - \frac{b}{3a}\right)^2 + c\left(x - \frac{b}{3a}\right) + d && (5.10)\\
&= a\left(x^3 - 3\frac{b}{3a}x^2 + 3(\frac{b}{3a})^2x - \frac{b^3}{27a^3}\right) + b\left(x^2 - \frac{2b}{3a}x + \frac{b^2}{9a^2}\right) + cx - \frac{bc}{3a} + d && (5.11)\\
&= ax^3 - bx^2 + \frac{b^2}{3a}x - \frac{b^3}{27a^2} && (5.12)\\
&\quad + bx^2 - \frac{2b^2}{3a}x + \frac{b^3}{9a^2} && (5.13)\\
&\quad\quad + cx - \frac{bc}{3a} + d && (5.14)\\
&= ax^3 + \frac{b^2}{3a}(1 - 2 + c)\,x + \frac{b^3}{9a^2}\left(1 - \frac{1}{3}\right) - \frac{bc}{3a} + d && (5.15)
\end{aligned}$$

The we move it in y direction by $-(\frac{2b^3}{27a^2} - \frac{bc}{3a} + d)$:

$$f_2(x) = ax^3 + \frac{b^2(c-1)}{3a}x \quad \text{and} \quad P_2 = (x_P + \frac{b}{3a}, y_P - (\frac{2b^3}{27a^2} - \frac{bc}{3a} + d))$$

Multiply everything by $\text{sgn}(a)$:

$$f_3(x) = \underbrace{|a|}_{=:\alpha}\, x^3 + \underbrace{\frac{b^2(c-1)}{3|a|}}_{=:\beta}\, x \quad \text{and} \quad P_2 = (x_P + \frac{b}{3a}, \text{sgn}(a)(y_P - \frac{2b^3}{27a^2} + \frac{bc}{3a} - d))$$

Now the problem seems to be much simpler. The function $\alpha x^3 + \beta x$ with $\alpha > 0$ is centrally symmetric to $(0, 0)$.

Und weiter?

5.1.3 Number of points with minimal distance

As this leads to a polynomial of degree 5 of which we have to find roots, there cannot be more than 5 solutions.

Can there be 3, 4 or even 5 solutions? Examples!
After looking at function graphs of cubic functions, I'm pretty sure that there cannot be 4 or 5 solutions, no matter how you chose the cubic function f and P.
I'm also pretty sure that there is no polynomial (no matter what degree) that has more than 3 solutions.

5.1.4 Interpolation and approximation

Quadratic spline interpolation

You could interpolate the cubic function by a quadratic spline.

Bisection method

TODO

Newtons method

One way to find roots of functions is Newtons method. It gives an iterative computation procedure that can converge quadratically if some conditions are met:

Theorem 7 (local quadratic convergence of Newton's method)
Let $D \subseteq \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^n \in C^2(\mathbb{R})$. Let $x^* \in D$ with $f(x^*) = 0$ and the Jaccobi-Matrix $f'(x^*)$ should not be invertable when evaluated at the root.

Then there is a sphere

$$K := K_\rho(x^*) = \{ x \in \mathbb{R}^n \mid \|x - x^*\|_\infty \leq \rho \} \subseteq D$$

such that x^* is the only root of f in K. Furthermore, the elements of the sequence

$$x_{n+1} = x_n - \frac{f'(x_n)}{f(x_n)}$$

are for every starting value $x_0 \in K$ again in K and

$$\lim_{n \to \infty} x_k = x^*$$

Also, there is a constant $C > 0$ such that

$$\|x^* - x_{n+1}\| = C\|x^* - x_n\|^2 \text{ for } n \in \mathbb{N}_0\|$$

The approach is extraordinary simple. You choose a starting value x_0 and compute

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

As soon as the values don't change much, you are close to a root. The problem of this approach is choosing a starting value that is close enough to the root. So we have to have a "good" initial guess.

Muller's method

Muller's method was first presented by David E. Muller in 1956.

Bisection method

The idea of the bisection method is the following:

Suppose you know a finite intervall $[a, b]$ in which you have exactly one root $r \in (a, b)$ with $f(r) = 0$.

Then you can half that interval:

$$[a, b] = \left[a, \frac{a+b}{2}\right] \cup \left[\frac{a+b}{2}, b\right]$$

Now three cases can occur:

Case 1 $f(\frac{a+b}{2}) = 0$: You have found the exact root.

Case 2 $\text{sgn}(a) = \text{sgn}(\frac{a+b}{2})$: Continue searching in $[\frac{a+b}{2}, b]$

Case 3 $\text{sgn}(b) = \text{sgn}(\frac{a+b}{2})$: Continue searching in $[a, \frac{a+b}{2}]$

Bairstow's method

Cite from Wikipedia: The algorithm first appeared in the appendix of the 1920 book "Applied Aerodynamics" by Leonard Bairstow. The algorithm finds the roots in complex conjugate pairs using only real arithmetic.

[...]

Find a source for the following!

A particular kind of instability is observed when the polynomial has odd degree and only one real root.

5.2 Defined on a closed interval $[a, b] \subseteq \mathbb{R}$

The point with minimum distance can be found by:

$$\arg\min_{x \in [a,b]} d_{P,f}(x) = \begin{cases} S_3(f, P) & \text{if } S_3(f, P) \cap [a, b] \neq \emptyset \\ TODO & \text{if } S_3(f, P) \cap [a, b] = \emptyset \end{cases}$$