Needham, Chapter 5, Exercise 10, p. 258.
G. Palmer, January 15, 2016

10. Consider the polynomial $P(z) = (z - a_1)(z - a_2) \ldots (z - a_n)$.

(i) Show that the critical points of P(z) are the solutions of

$$\frac{1}{z-a_1} + \frac{1}{z-a_2} + \ldots + \frac{1}{z-a_n} = 0.$$

The critical points of P(z) are the points for which P'(z) = 0. So we calculate P'(z) using Leibnitz's rule (fg)' = f'g + fg' recursively.

let $f_i = z - a_i$, where i ∋ n

Then $P(z)' = (f_1 f_2 \ldots f_n)'$ and $(f_n)' = 1$ for all n.

$$P(z)' = (f_1)'(f_2 f_3 \ldots f_n) + f_1(f_2 f_3 \ldots f_n)'$$

$$= f_2 f_3 \ldots f_n + f_1((f_2)'(f_3 f_4 \ldots f_n) + f_2(f_3 f_4 \ldots f_n)')$$

$$= f_2 f_3 \ldots f_n + f_1 f_3 f_4 \ldots f_n + f_1 f_2(f_3 f_4 \ldots f_n)'$$

$$= f_2 f_3 \ldots f_n + f_1 f_3 f_4 \ldots f_n + f_1 f_2((f_3)'(f_4 \ldots f_n) + f_3(f_4 \ldots f_n)')$$

$$= f_2 f_3 \ldots f_n + f_1 f_3 f_4 \ldots f_n + f_1 f_2 f_4 \ldots f_n + f_1 f_2 f_3 f_5 \ldots f_n + \ldots + f_1 f_2 \ldots f_{n-2}(f_{n-1} f_n)'$$

$$= f_2 f_3 \ldots f_n + f_1 f_3 f_4 \ldots f_n + f_1 f_2 f_4 \ldots f_n + f_1 f_2 f_3 f_5 \ldots f_n + \ldots + f_1 f_2 \ldots f_{n-2}(f_{n-1})' f_n + f_{n-1}(f_n)')$$

$$= f_2 f_3 \ldots f_n + f_1 f_3 f_4 \ldots f_n + f_1 f_2 f_4 \ldots f_n + f_1 f_2 f_3 f_5 \ldots f_n + \ldots + f_1 f_2 \ldots f_{n-2} f_n + f_1 f_2 \ldots f_{n-2} f_{n-1}$$

It appears that each summand loses one f_i term from the sequence of P(z). The index of the missing term is the same as the index of the summand from which it is missing. For example, f_1 does not appear the first summand and f_n does not appear in the last. The number of summands equals the degree of P(z). Rewrite:

$$P(z) = \prod_{1=i}^{n} f_i(z)$$

$$P(z)' = \sum_{1=j}^{n} P(z) / f_j$$

$$= P(z) \sum_{1=j}^{n} \frac{1}{f_j}$$

Since the critical points of P(z) are the points for which P(z)' = 0, we take

$$P(z) \sum_{1=j}^{n} \frac{1}{f_j} = 0 \qquad (1)$$

and divide both sides by P(z).

$$\frac{P(z)}{P(z)} \sum_{1=j}^{n} \frac{1}{f_j} = \sum_{1=j}^{n} \frac{1}{f_j} = 0$$

Substitute the $z - a_i$ terms back into the quotients for each f_j to get the equation in (i)

$$\sum_{1=j}^{n} \frac{1}{f_j} = \sum_{1=j}^{n} \frac{1}{z-a_j} = \frac{1}{z-a_1} + \frac{1}{z-a_2} + \ldots + \frac{1}{z-a_n} = 0$$

We have calculated P(z)'/P(z) to obtain a simplified but still valid equation in z where P(z)' = 0, showing that the solutions for z are the critical points of P(z).

(ii) Let K be a circle with centre p. By considering the conjugate of the equation in (i), deduce that p is a critical point if and only if it is the centre of mass of the inverted points $\mathcal{I}_K(a_j)$.

This would mean that the centre of mass of the inverted points is a solution to the equation in (i) as well as a critical point.

The conjugate of the equation in (i):

$$\overline{\sum_{j=1}^{n} \frac{1}{f_j}} = \sum_{j=1}^{n} \overline{\frac{1}{f_j}} = \sum_{1=j}^{n} \frac{1}{\overline{f_j}}$$

$$= \frac{1}{\overline{z-a_1}} + \frac{1}{\overline{z-a_2}} + \ldots + \frac{1}{\overline{z-a_n}}$$

$$= \frac{1}{\overline{z}-\overline{a}_1} + \frac{1}{\overline{z}-\overline{a}_2} + \ldots + \frac{1}{\overline{z}-\overline{a}_n} = 0 \qquad (2)$$

Note that the conjugate of P'(z) is equivalent to P'(z).

$$\mathcal{I}_K(a_j) = \frac{R^2}{\overline{a}_j - \overline{p}} + \text{p} \quad \text{(for any K centred at p)}$$

We assume equal masses for all of the inverted points.

$$\text{Centroid}(\mathcal{I}_K(a_j)) = \frac{1}{n} \sum_{j=1}^{n} \left(\frac{R^2}{\overline{a}_j - \overline{p}} + p \right) \qquad \text{(p. 104)}$$

$$= \frac{1}{n} \sum_{j=1}^{n} \frac{R^2}{\overline{a}_j - \overline{p}} + \frac{\text{np}}{p}$$

$$= \frac{1}{n} \sum_{j=1}^{n} \frac{R^2}{\overline{a}_j - \overline{p}} + p \qquad (3)$$

If the centroid equals p, then $\frac{1}{n}\sum_{j=1}^{n}\left(\frac{R^2}{\bar{a}_j-\bar{p}}\right) = 0$ and, after multiplying both sides by $\frac{n}{R^2}$,

$$\sum_{j=1}^{n}\left(\frac{1}{\bar{a}_j-\bar{p}_j}\right) = \frac{1}{\bar{z}-\bar{a}_1} + \frac{1}{\bar{z}-\bar{a}_2} +\ldots+ \frac{1}{\bar{z}-\bar{a}_n} = 0$$

Since -0 = 0,

$$-\sum_{j=1}^{n}\left(\frac{1}{\bar{a}_j-\bar{p}}\right) = \frac{1}{\bar{z}-\bar{a}_1} + \frac{1}{\bar{z}-\bar{a}_2} +\ldots+ \frac{1}{\bar{z}-\bar{a}_n}$$

$$\sum_{j=1}^{n}\left(\frac{1}{\bar{p}-\bar{a}_j}\right) = \frac{1}{\bar{z}-\bar{a}_1} + \frac{1}{\bar{z}-\bar{a}_2} +\ldots+ \frac{1}{\bar{z}-\bar{a}_n} \qquad (4)$$

Conjugating both sides, we get

$$\sum_{j=1}^{n}\left(\frac{1}{p-a_j}\right) = \frac{1}{z-a_1} + \frac{1}{z-a_2} +\ldots+ \frac{1}{z-a_n} = 0 \qquad (5)$$

This shows that p is a solution to P'(z) when P'(z) = 0, so p is a critical point when p is the centroid of the inverted points $\mathcal{I}_K(a_j)$ for all j in n. If p were not a centroid for $\mathcal{I}_K(a_j)$, then $\sum_{j=1}^{n}\left(\frac{1}{\bar{p}-\bar{a}_j}\right) \neq 0$, from which it follows that $\sum_{j=1}^{n}\left(\frac{1}{p-a_j}\right) \neq 0$, and $\sum_{j=1}^{n}\left(\frac{1}{p-a_j}\right) \neq \frac{1}{z-a_1} + \frac{1}{z-a_2} +\ldots+ \frac{1}{z-a_n}$, so p would not be a solution to P(z) when P(z)' = 0. Therefore, p is a critical point if and only if it is the centre of mass of the inverted points $\mathcal{I}_K(a_j)$.

(iii) Show that the equation in (i) is equivalent to

$$\frac{z-a_1}{|z-a_1|^2} + \frac{z-a_2}{|z-a_2|^2} + \ldots + \frac{z-a_n}{|z-a_n|^2} = 0$$

and by interpreting the LHS as a (positively) weighted sum of the vectors from z to the roots of P(z), deduce Lucas' Theorem: The critical points of a polynomial in ℂ must all lie within the convex hull of its zeroes. This is a complex generalization of Rolle's Theorem in ordinary calculus. [Hint: Use the fact that (32) on page 104 is still valid even if the masses are not equal.]

The equation shows that if the LHS = 0, the sum of the vectors from z to the roots of P(z) is a centroid for f_i. But the zeroes of P(z) are the a_i, not the f_j. Apparently, we must find an equation containing both the critical points and the a_i. We first try to satisfy the first part of the question.

From (2), $\sum_{j=1}^{n}\frac{1}{\bar{f}_j} = \sum_{j=1}^{n}\frac{1}{f_j} = 0$. Multiply each term of $\sum_{j=1}^{n}\frac{1}{\bar{f}_j}$ by 1 in the form of $\frac{f_j}{f_j}$.

$$\sum_{j=1}^{n}\left(\frac{1}{\bar{f}_j}\frac{f_j}{f_j}\right) = \sum_{j=1}^{n}\frac{f_j}{|f_j|^2} = \frac{z-a_1}{|z-a_1|^2} + \frac{z-a_2}{|z-a_2|^2} + \ldots + \frac{z-a_n}{|z-a_n|^2} = 0 \qquad (6)$$

Deduce Lucas' Theorem. The real factors $1/|f_n|^2$ may be taken as equivalent to the mass by the first equation of Section VIII.1 The Centroid, p. 103, so we can write (6) as

$\frac{\sum_{j=1}^{n} m_j f_j}{\sum_{j=1}^{n} m_j} = 0$, where $m_j = \frac{1}{|f_j|^2}$

Multiply both sides by $\sum_{j=1}^{n} m_j$.

$$\sum_{j=1}^{n} m_j f_j = 0$$

$$= \sum_{j=1}^{n} m_j (z - a_j)$$

$$= \sum_{j=1}^{n} m_j z - \sum_{j=1}^{n} m_j a_j$$

$$z \sum_{j=1}^{n} m_j = \sum_{j=1}^{n} m_j a_j$$

$$z = \frac{\sum_{j=1}^{n} m_j a_j}{\sum_{j=1}^{n} m_j} \qquad (7)$$

This is an expression for the centroid of a set of particles a_j of masses m_j. Any z that is a solution to (7) is a critical point of P(z) because it also satisfies equation in (i); it is a point for which P'(z) = 0. By the hint, 32 is still valid even if the masses are unequal, so the centroid z, which is also a critical point, must be in the interior of the convex hull H of the zeroes of P(z), which are the points a_j. There will be a different set of masses at the zeros a_j, for each critical point.

Please note that the following is not an answer to the question. It is just my own thinking about the relationships in the problem. In essence, The critical points (solutions to P') are also centroids of the zeroes of P, so the critical points must lie within the hull . We break this down:

The centroid must lie in the interior of the complex hull H. (32)

z is a centroid with it's masses placed at a_j, as can be seen by comparing the equation for z in (iii) to the equation for Z at the top of p. 103.

The complex hull in question is the smallest polygon constructed on a_j, which are the zeros of P. That the a_j compose the complex hull can be seen by comparing the equation for z in (iii) to the first equation for Z on p. 103 and comparing that to Figure[38].

Therefore, z lies within the complex hull of the zeros.

z represents all the solutions to (iii), which is the weighted version of P'(z).

By (i), the solutions to P'(z) are the critical points of P.

Therefore, the critical points of P must lie within the complex hull.

Thinking about the visual aspects of the problem, by (i), the solutions to P' are critical points. By (iii), the solutions are centroids with masses of particles placed at the zeroes. Then, since these solutions are

solutions to the same equation P' = 0, these centroids must be critical points, and since the centroids obviously in most cases lie in the interior of the complex hull and in fact must lie within the hull by (32), we can see that the same is true for the critical points.

We might then restate the problem: Show that the critical points (found in (i)) lie within the hull by showing that they are centroids of the zeros. Show this by showing that the centroids of the zeroes are solutions to P'. This can be done by deriving the equation for the centroids of the zeroes from the equation for P'.

The plan for the Figures:

Figure 1.
1. plot a random set of dots a_j in black
2. plot the centroid of a_j in black with a circle around it using the centroid formula $\frac{\sum_{j=1}^{n} m_j a_j}{\sum_{j=1}^{n} m_j}$
3. plot unit circle K as dashed gray circle centred on p, the centroid of a_j
4. plot centroid of a_j as black dot surrounded by black circle
5. plot $\mathcal{I}_K(a_j)$, the inverses of a_j in blue by inverting in circle K centred on p, the centroid of a_j
6. plot the centroid of $\mathcal{I}_K(a_j)$ in blue with a circle around it using the centroid formula w/o mass $\frac{1}{n}\sum_{j=1}^{n} \mathcal{I}_K(a_j)$

Figure 2.
1. plot a small gray circle of points around each centroid, C_{a_j} and $C_{\mathcal{I}_K(a_j)}$
2. apply P(z) to the points in each circle and plot the results

In Figure 1, the 5 black dots represent a_j in the problem. They are a random set with both real and imaginary parts ranging from -.5 to .5. The dot inside a circle is their centroid. The centroid is used for p, which is the centre of the circle of inversion K, which is a unit circle. The blue dots are the $I_k(a_j)$. The blue dot inside the blue circle is the centroid of the inverted dots. The centroids of a_j and $I_k(a_j)$ are not equal.

In Figure 2 we plot the two centroids again, and we plot two small gray circles, one around each centroid. When we plot $P(z)$ on the points of the small gray circle around C_{a_j}, which is the centroid of a_j, we do find that the circle crushes down to a point (not shown). When we scale the crushed points up by a factor of 100, we find that image of the circle is irregular (Orange points). For reasons unknown, we have been unable to plot $P(z)$ on the points in the small gray circle around $\mathcal{I}_K(a_j)$. On some random results, the points of P(z) on the circle appear to be plottable, but they do not plot.

centroid a_j: −0.0904667 + 0.0357862 I

centroid $\mathcal{I}_K(a_j)$: −0.170839 + 0.0611433 I

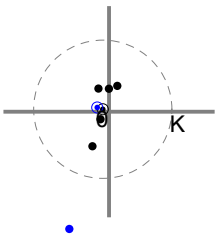

Figure 1. Critical points and centroids
of P(z) = (z−a_1)(z−a_2)...(z−a_n)
a_j: Black, $\mathcal{I}_K(a_j)$: Blue

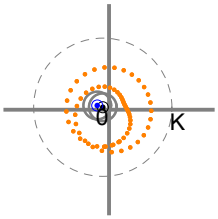

Figure 2. Critical points and centroids
of P(z) = (z–a_1)(z–a_2)...(z–a_n)
Centroid(a_j): Black, Centroid($\mathcal{I}_K(a_j)$): Blue
P(z) on small gray circle around Centroid(a_j):
Orange and Scaled by 100