NOTES AND EXERCISES ON CONSTRUCTIBILITY

A collection $\mathcal{C}$ of points, lines, and circles in the euclidean plane is *constructibly closed* if it satisfies the following conditions:

(1) $\mathcal{C}$ contains both 0 and 1.

(2) (Euclid's first and second postulates.) A line which contains two points of $\mathcal{C}$ is in $\mathcal{C}$.

(3) (Euclid's third postulate.) A circle which contains a point of $\mathcal{C}$, and whose center is in $\mathcal{C}$, is in $\mathcal{C}$.

(4) If a point P lies on more than one member of $\mathcal{C}$, then $P \in \mathcal{C}$.

Let C be any collection of points, lines, and circles, containing 0 and 1. The members of C are called C-points, resp. C-lines, resp. C-circles. Set $C_0 := C$ and for $n \geq 0$, proceeding recursively, set

$$\begin{aligned} C_{n+1} := C_n &\cup \{\text{ all lines containing two } C_n\text{-points}\} \\ &\cup \{\text{ all circles with center in } C_n \text{ and containing a } C_n\text{-point}\} \\ &\cup \{\text{ all points lying on more than one member of } C_n\}. \end{aligned}$$

It is straightforward to check that:

Proposition 1. *$C_\infty := \bigcup_{n=0}^{\infty} C_n$ is the smallest constructibly closed collection containing C.*

We say that a point, or line, or circle, is *constructible from C* if it is a member of C_∞. This is a precise formulation of what it means to construct something with straightedge and compass, starting from the initial data C.

The next thing is to algebraicize the foregoing considerations. In what follows, we identify the euclidean plane with the set of complex numbers $\mathbb{C}$; and all fields are assumed to be subfields of $\mathbb{C}$.

A field is *constructibly closed* if it is closed under square roots and under complex conjugation.

Let $\mathcal{C}$ be a constructibly closed collection of points, lines, and circles.

Using simple straightedge-and-compass-constructions, one sees that if P is a $\mathcal{C}$-point and L is a $\mathcal{C}$-line then the line perpendicular to L and passing through P is a $\mathcal{C}$-line. Hence $a + ib$ $(a, b \in \mathbb{R})$ is a $\mathcal{C}$-point if and only if so are a and b.

We showed in class that:

$(*)$ The $\mathcal{C}$-points form a constructibly closed field.

It follows, since 0 and 1 are $\mathcal{C}$-points, that $a + bi$ is a $\mathcal{C}$-point for all $a, b \in \mathbb{Q}$, and that every line $aX + bY = c$ with $a, b, c \in \mathbb{Q}$ is a $\mathcal{C}$-line.

The following exercises lead to the conclusion that $(*)$ actually gives a one-one correspondence between constructibly closed collections and constructibly closed fields.

(a) Prove that a $\mathcal{C}$-line contains two $\mathcal{C}$-points.

(b) Prove that on any $\mathcal{C}$-circle there lie three (in fact, infinitely many) $\mathcal{C}$-points.

(c) Use elementary straightedge-and-compass-constructions to show that the center of a $\mathcal{C}$-circle is a $\mathcal{C}$-point.

(d) Explain why $\mathcal{C}$ is determined by the set of its points. (Use (a), (b) and (c).)

(e) Given real numbers a, b, c, and d, not all 0, show that the set of points $(x, y) \in \mathbb{R}^2$ satisfying the relation

$$a(x^2 + y^2) + bx + cy + d = 0$$

are all the points in a set of one of the following types:

(i) The empty set;

(ii) A set consisting of one point;

(iii) A line;

(iv) A circle.

Show conversely that any set of one of the preceding types is defined by a relation of the given form.

(f) In (e), let e be any one of a, b, c, d which is not zero. In cases (ii), (iii), and (iv), show that the set is in $\mathcal{C}$ if and only if the ratios a/e, b/e, c/e, and d/e are all in $\mathcal{C}$.

(g) Let F be a constructibly closed field. Prove that the points of F together with all the lines joining pairs of points in F together with all the circles whose center is in F and which pass through at least one point of F form a constructibly closed collection.

Theorem 1. *There is a* one-one correspondence *between constructibly closed collections and constructibly closed fields, obtained by associating to each constructibly closed collection its set of points.*

Proof. This follows from $(*)$, (d) and (g).

Definition. *The field of definition* of a nonempty set as in (f) above is the field $\mathbb{Q}(a/e, b/e, c/e, d/e)$. The field of definition of a collection C as above is the field generated by the fields of definition of all the members of C.

Remarks. (i) The field of definition of a point $a + ib$ is $\mathbb{Q}(a, b)$.

(ii) If C is constructibly closed, then its field of definition is just the field consisting of all of its points.

(iii) From (f) above, it follows that the points in the field of definition of C are constructible from C.

Corollary. *Let C be as above, with field of definition F. Then the points of C_∞ form the smallest constructibly closed field containing F.*

At this point, we have translated the geometric constructibility problem—that is, given C, determine whether or not a point, line or circle lies in C_∞—into a purely algebraic one, namely, *given a field F, determine whether or not a point P lies in the smallest constructibly closed field $\mathcal{E}_F$ containing F.*

We'll say that a point P is *constructible from F* if $P \in \mathcal{E}_F$.

Theorem 2. *Let F be a field that is closed under complex conjugation. A point P is constructible from F if and only if it "sits in a quadratic tower based on F," i.e., there exists a sequence of fields*

$$F = F_0 \subset F_1 \subset \cdots \subset F_n$$

such that $[F_i : F_{i-1}] = 2$ for all $i = 1, 2, \ldots, n$ and such that $P \in F_n$.

Corollary. *If P is constructible from F then $[F(P) : F]$ is a power of 2.*

Remarks. (iv) P sits in a quadratic tower if and only if

$$P \in F[x_1, x_2, \ldots, x_n]$$

where $x_i^2 \in F[x_1, x_2, \ldots, x_{i-1}]$ for $1 \leq i \leq n$. That's because any degree-2 extension of a field K is obtained by adjoining the root of a quadratic equation, hence (as long as K has characteristic $\neq 2$) by adjoining the square root of some element of K.

Roughly speaking, for P to sit in a quadratic tower means that P can be constructed starting with an element of F and applying addition, subtraction, multiplication, division, and square root a finite number of times. (And, recall, each of these operations can be done with straightedge and compass.)

(v) Let z be a root of the polynomial $X^4+X+1=0$, and set $F=\mathbb{Q}[z]$. Then the complex conjugate $\bar{z}$ is constructible from F (since $\mathcal{E}_F$ is closed under complex conjugation), and it is a root of the polynomial $(X^4+X+1)/(X-z) \in F[X]$. It can be shown that this polynomial is irreducible over F, and so $[F(\bar{z}):F]=3$. Thus $\bar{z}$ cannot sit in any quadratic tower based on F.

This doesn't contradict the Theorem, because F is not closed under complex conjugation.

Proof of Theorem 2. ($\Rightarrow$) Since the set of points in $\mathcal{E}_F$ is closed under square roots, Remark (iv) shows that every F_n as above is contained in that set of points. Thus any $P \in F_n$ is constructible from F.

($\Leftarrow$) It suffices to show that the set G of all points which sit in some quadratic tower form a constructibly closed field—since that forces $\mathcal{E}_F$ (the smallest constructibly closed field containing F) to be contained in G.

First, if P sits in a quadratic tower, then so does its complex conjugate, as one can see by applying conjugation to the entire tower. (This is where we use that F is closed under conjugation.) Thus G is closed under conjugation.

Next, suppose that P and Q are in G, so that as in Remark (iv),

$$P \in F[x_1, x_2, \ldots, x_n], \qquad Q \in F[y_1, y_2, \ldots, y_m].$$

Then we have the quadratic tower obtained from F by first adjoining the P's one at a time, and then adjoining the y's one at a time. Both P and Q are in the top level of this tower, whence so is $P+Q$, so that $P+Q \in G$. Similarly PQ and (if $Q \neq 0$) P/Q are members of G. This shows that G is a field.

If P sits in some tower $F=F_0 \subset \cdots \subset F_n$, then $\sqrt{P}$ sits in the tower $F=F_0 \subset \cdots \subset F_n \subset F_n(\sqrt{P})$. Thus G is closed under square roots. $\square$

Finally, here is a variant of Theorem 1, involving only real numbers:

Theorem 1′. *There is a* one-one correspondence *between constructibly closed collections and subfields of $\mathbb{R}$ closed under square roots, obtained by associating to each constructibly closed collection the least field containing the (real) coordinates of all its points.*

Proof. One could imitate the proof of Theorem 1, *mutatis mutandis.* Or, if Theorem 1 is assumed then one can use the easily-shown fact that by associating to each constructibly closed field F the field $G := F \cap \mathbb{R}$, one gets a one-one correspondence between constructible fields and subfields of $\mathbb{R}$ closed under square roots. (Show that $a+ib \in F \iff a \in G$ and $b \in G$, and conclude that $F=G[i]$.)