Algorithms and Geometric Constructions

Vladimir Uspenskiy, Alexander Shen

Algorithms and Geometric Constructions

Vladimir Uspenskiy[1] and Alexander Shen[2]

[1] Moscow State Lomonosov University
[2] LIRMM CNRS and University of Montpellier. On leave from IITP RAS.[*]

Abstract. It is well known that several classical geometry problems (e.g., angle trisection) are unsolvable by compass and straightedge constructions. But what kind of object is proven to be non-existing by usual arguments? These arguments refer to an intuitive idea of a geometric construction as a special kind of an "algorithm" using restricted means (straightedge and/or compass). However, the formalization is not obvious, and different descriptions existing in the literature are far from being complete and clear. We discuss the history of this notion and a possible definition in terms of a simple game.

1   Introduction

The notion of an algorithm as an intuitively clear notion that precedes any formalization, has a rather short history. The first examples of what we now call algorithms were given already by Euclid and al-Khwârizmî. But the general idea of an algorithm seems to appear only in 1912 when Borel considered "les calculus qui peuvent être réellement effectués"[3] and emphasized: "Je laisse intentionnellement de côté le plus ou moins grande longeur pratique des opérations; l'essentiel est que chaque de ces opérations soit exécutable en un temps fini, par une méthode sûre et sans ambiguïte"[4] [4, p. 162]. The formal definition of a representative class of algorithms was given in 1930s (in the classical works of Gödel, Church, Kleene, Turing, Post and others); the famous Church–Turing thesis claims that the class of algorithms provided by these formal definitions is representative.

In this paper we look at the history of another related notion: the notion of a *geometric construction*. One may consider geometric constructions as a special type of algorithms that deal with geometric objects. Euclid provided many examples of geometric constructions by compass and straightedge (ruler); later these constructions became a standard topic for high school geometry exercises. Several classical problems (angle trisection, doubling the square, squaring the circle) were posed and remained unsolved since ancient times (though solutions

---

[3] the computations that can be really performed
[4] I intentionally put aside the question of bigger or smaller practical length of the operation; it is important only that each of the operations can be performed in a finite time by a clear and unambiguous method".

that involve more advanced instruments than compass and straightedge were suggested). These problems were proved to be unsolvable in XIX century. One would expect that the proof of unsolvability assumes as a prerequisite a rigorously defined notion of a "solution" that does not exist. (Recall that the first undecidability proofs could appear only after an exact definition of an algorithm was given.)

However, historically this was not the case and the impossibility proofs appeared without an exact definition of a "geometric construction". These proofs used the algebraic approach: For example, to show that the cube cannot be doubled, one proves that $\sqrt[3]{2}$ cannot be obtained from rationals by arithmetic operations and square roots. The reduction from a geometric question to an algebraic one looks quite obvious and was omitted by Wantzel who first proved the impossibility of angle trisection and cube doubling. As he wrote in [22], "pour reconnaitre si la construction d'un problème de Géometrie peut s'effectuer avec la règle et le compas, if faut chercher s'il est possible de faire dépendre les racines de l'equation à laquelle il conduit de celles d'un système d'équations du second degré".[5] This is said in the first paragraph of the paper and then he considers only the algebraic question.

Several other interesting results were obtained in XIX century. It was shown that all constructions by compass and straightedge can be performed using the compass only (the *Mohr–Mascheroni theorem*) if we agree that a line is represented by a pair of points on this line. Another famous result from XIX century, the *Poncelet–Steiner theorem*, says that if a circle with its center is given, then the use of compass can be avoided, straightedge is enough. Other sets of tools were also considered, see, e.g., [3,14,9].

Later geometric construction became a popular topic of recreational mathematics (see, e.g., [15,6,13,10]). In most of the expositions the general notion of a geometric construction is still taken as granted, without a formal definition, even in the nonexistence proofs (e.g., when explaining Hilbert's proof that the center of a circle cannot be found using only a straightedge [15,6,10]; see below Section 6 about problems with this argument). Sometimes a definition for some restricted class of geometric construction is given (see, e.g., [18]). In [13] an attempt to provide a formal definition is made, still it remains ambiguous with respect to the use of "arbitrary points" (see Section 4). Baston and Bostock [2] observe that the intuitive idea of a "geometric construction" has no adequate formal definition and discuss several examples but do not attempt to give a formal definition that is close to the intuitive notion. It seems that even today people still consider the intuitive notion of a "geometric construction algorithm" as clear enough to be used without formal definition (cf. [1], especially the first `arxiv` version).

In Section 2 we consider a naïve approach that identifies constructible points with the so-called "derivable" points. Then in Sections 3 and 4 we explain why

---

[5] To find out whether a geometric problem can be solved by straightedge and compass construction, one should find whether it is possible to reduce the task of finding the roots of the corresponding equation to a system of equations of second degree.

this approach contradicts our intuition. In Section 5 we suggest a more suitable definition, and finally in Section 6 we note that the absence of formal definitions has led to incorrect proofs.

2   Derivable points and straight-line programs

At first it seems that the definition of a geometric construction is straightforward. We have three classes of geometric objects: points, lines and circles. Then we consider some operations that can be performed on these objects. We need to obtain some object (the goal of our construction) applying the allowed operations to given objects. As Tao [18] puts it,

> Formally, one can set up the problem as follows. Define a configuration to be a finite collection $\mathcal{C}$ of points, lines, and circles in the Euclidean plane. Define a construction step to be one of the following operations to enlarge the collection $\mathcal{C}$:
>
> – (Straightedge) Given two distinct points A, B in $\mathcal{C}$, form the line $\overline{AB}$ that connects A and B, and add it to $\mathcal{C}$.
> – (Compass) Given two distinct points A, B in $\mathcal{C}$, and given a third point O in $\mathcal{C}$ (which may or may not equal A or B), form the circle with centre O and radius equal to the length $|AB|$ of the line segment joining A and B, and add it to $\mathcal{C}$.
> – (Intersection) Given two distinct curves γ, γ' in $\mathcal{C}$ (thus γ is either a line or a circle in $\mathcal{C}$, and similarly for γ'), select a point P that is common to both γ and γ' (there are at most two such points), and add it to $\mathcal{C}$.
>
> We say that a point, line, or circle is constructible by straightedge and compass from a configuration $\mathcal{C}$ if it can be obtained from $\mathcal{C}$ after applying a finite number of construction steps.

We can even try to define the geometric construction algorithm as a straight-line program, a sequence of assignments whose left-hand side is a fresh variable and the right-hand side contains the name of the allowed operation and the names of objects to which this operation is applied.

Baston and Bostock [2] use the name "derivable" for objects that can be obtained in this way starting from given objects. In other word, starting with some set of given objects, they consider its closure, i.e., the minimal set of objects that contains the given ones and is closed under allowed operations. The objects that belong to this closure are called *derivable* from the given ones. In these terms, the impossibility of trisecting the angle with the compass and the straightedge can be states as follows: *for some points A, B, C the trisectors of the angle BAC are not derivable from* $\{A, B, C\}$.

Baston and Bostock note that the intuitive notion of a "constructible" point (that they intentionally leave without any definition) may differ from the formal notion of a derivable point in both directions. We discuss the differences in the following sections.

3   Uniformity and tests

There are some problems with this approach. First of them is its "non-uniformity". Asking a high school student to construct, say, a center of an inscribed circle of a triangle ABC, we expect the answer to be a construction that works *for all triangles*, not just the proof that this center is always derivable from A, B, and C. The naïve approach would be to ask for a straight-line program that computes this center starting from A, B, and C. However, an obvious problem arises: the operation of choosing an intersection point of two curves is non-deterministic (we need to choose one of two intersection points). We may guarantee only that *some* run of the program produces the required object, or guarantee that the required object is among the objects computed by this program. This is a common situation for classical constructions. For example, the standard construction of the centre of the incircle of a triangle can also produce centres of excircles (the circles outside the triangle that touch one of its sides and the extensions of two other sides).

The non-deterministic nature of the operations was mentioned by different authors. Bieberbach [3] says that the constructions should be performed in the "oriented plane" (not giving any definitions). Tietze [19,20,21] notes that some objects can be constructed but only in a non-deterministic way, again without giving definition of these notions.

One could give up and consider the non-uniform setting only. As Manin [13, p. 209] puts it, "we ignore how to choose the required point from the set of points obtained by the construction". Another approach is to replace straight-line programs by decision trees where tests appear as internal nodes. Still none of these two approaches (decision trees or non-deterministic choice) is enough to save some classical constructions in a uniform setting as observed by Baston and Bostock [2, p. 1020]. They noted that the construction from Mohr–Mascheroni theorem allows us to construct the intersection point of two intersecting lines AB and CD (given A, B, C, D) using only a compass. Each use of compass increases the diameter of the current configuration at most by $O(1)$-factor, and the intersection point can be arbitrarily far even if A, B, C, D are close to each other, so there could be no *a priori* bound on the number of steps. (The necessity of an iterative process in the Mohr–Mascheroni theorem was earlier mentioned in another form by Dono Kijne [11, ch. VIII, p. 99]; he noted that this result depends on Archimedes' axiom.)

To save the Mohr–Mascheroni construction, one may consider programs that allow loops. This was suggested, e.g., by Engeler [7]. Here we should specify what kind of data structures are allowed (e.g., whether we allow dynamic arrays of geometric objects or not). In this way we encounter another problem, at least if we consider straightedge-only constructions on the *rational* plane $\mathbb{Q}^2$ and allow using tests and unbounded number of steps/objects. Baston and Bostock [2] observed that having four different points $A, B, C, D \in \mathbb{Q}^2$ in a general position (no three points lie on a line, no two connecting lines are parallel), we can enumerate all (rational) points and therefore all rational lines. Then we can wait until a line parallel to AB appears (we assume that we may test whether two

given lines intersect or are parallel) and then use this parallel line to find the midpoint of AB. This construction does not look like a intuitively valid geometric construction (and contradicts the belief that one cannot construct the midpoint using only a straightedge, see [2] for details).

4   Arbitrary points

Let us now consider the other (and probably more serious) reason why the notion of a derivable object differs from the intuitive notion of a constructible object. Recall the statement about angle trisection as stated by Tao [18]: for some triangle ABC the trisectors of angle BAC are not derivable from $\{A, B, C\}$. (Tao uses the word "constructible", but we keep this name for the intuitive notion, following [2].) Tao interprets this statement as the impossibility of angle trisection with a compass and straightedge, and for a good reason.

On the other hand, the center of a circle is not derivable from the circle itself, for the obvious reason that no operation can be applied to enlarge the collection that consists only of the circle. Should we then say that the center of a given circle cannot be constructed by straightedge and compass? Probably not, since such a construction is well known from the high school. Similar situation happens with the construction of a bisector of a given angle (a configuration consisting of two lines and their intersection point).

Looking at the corresponding standard constructions, we notice that they involve another type of steps, "choosing an arbitrary point" (on the circle or elsewhere). But we cannot just add the operation "add an arbitrary point" to the list of allowed operations, since all points would become derivable. So what are the "arbitrary points" that we are allowed to add? Bieberbach [3, p. 21] speaks about "Punkte, über die keine Angaben affiner oder metrischer Art gemacht sind"[6] and calls them "willkürliche Punkte"-– but this hardly can be considered as a formal definition.

Tietze [21] notes only that "the role of arbitrary elements is not so simple as it is sometimes thought". Baston and Bostock [2] explain the role of arbitrary elements, but say only that "the distinction between constructibility and derivability arising from the use of arbitrary points is not very complex" and "we will not pursue a more detailed analysis in this direction"; they refer to [12] for an "elementary approach", but this book also does not give any clear definition. Probably the most detailed explanation of the role of arbitrary points is provided by Manin [13], but he still defines the construction as a "finite sequence of steps" (including the "arbitrary choices") and says that point is constructible if there exists a construction that includes this point "for all possible intermediate arbitrary choices"; this definition, if understood literally, makes no sense since different choices leads to different constructions. Schreiber [16] tries to define the use of arbitrary points in a logical framework, but his exposition is also far from being clear.

----

[6] points for which we do not have affine or metric information

How can we modify the definitions to make them rigorous? One of the possibilities is to consider the construction as a strategy in some game with explicitly defined rules. We discuss this approach in the next section.

5   Game definition

The natural interpretation of the "arbitrary choice" is that the choice is made by an adversary. In other words, we consider a game with two players, Alice and Bob. We start with the non-uniform version of this game.

Let E be some finite set of geometric objects (points, lines, and circles). To define which objects x are constructible starting from E, consider the following full information game. The position of the game is a finite set of geometric objects. The initial position is E. During the game, Alice and Bob alternate. Alice makes some requests, and Bob fulfills these requests by adding some elements to the current position. Alice wins the game when x becomes the element of the current position. The number of moves is unbounded, so it is possible that the game is infinite (if x never appears in the current position); in this case Alice does not win the game.

Here are possible request types.

– Alice may ask Bob to add to the current position some straight line that goes through two different points from the current position.
– Alice may ask Bob to add to the current position a circle with center A and radius BC, if A, B, C are points from the current position.
– Alice may ask Bob to add to the current position one or two points that form the intersection of two different objects (lines or circles) that already belong to the current position.

If we stop here, we get exactly the notion of derivable points, though in a strange form of a "game" where Bob has no choice. To take the "arbitrary" points into account, we should add one more operation:

– Alice specifies an open subset of the plane (say, an open circle), and Bob adds some point of this subset to the current position.

The point x is *constructible* from E if Alice has a winning strategy in this game.

Let us comment on the last operation.

(1) Note that Alice cannot (directly) force Bob to choose some point on a line or on a circle, and this is often needed in the standard geometric constructions. But this is inessential since Alice can achieve this goal in several steps. First she asks to add points on both sides of the line or circle (near the point that she wants to be added), then asks to connect these points by a line, and then asks to add the intersection point of this new line and the original one.

(2) On the other hand, according to our rules, Alice can guarantee that the new point is in a small neighborhood of the desired position. A weaker (for Alice) option would be to allow her to select a connected component of the complement of the union of all objects in the current position. Then Bob should add some point of this component to the current position.

Proposition 1. *This restriction does not change the notion of a constructible point.*

Proof. Idea: Using the weaker option, Alice may force Bob to put enough points to make the set of derivable points dense, and then use the first three options to get a point in an arbitrary open set.

Let us explain the details. First, she asks for an arbitrary point A, then for a point B that differs from A, then for line AB, then for a point C outside line AB (thus having the triangle ABC), then for sides of this triangle, and then for a point D inside the triangle. (All this is allowed in the restricted version.)
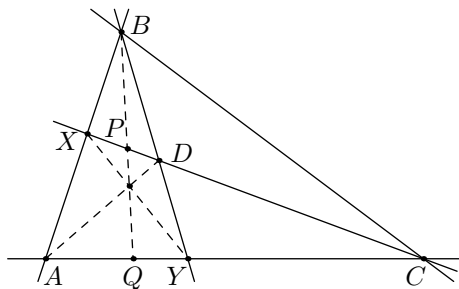
Now the points P and Q obtained as shown are derivable (after the projective transformation that moves B and C to infinity, the points P and Q become the midpoints of XD and AY). Repeating this construction, we get a dense set of derivable points on intervals XD and AY, then the dense set of derivable points in the quadrangle $AXDY$ and then in the entire plane.

Now, instead of asking Bob for a point in some open set U, Alice may force him to include one of the derivable points (from the dense set discussed above) that is in U.

This definition of constructibility turns out to be equivalent to the negative definition suggested by Akopyan and Fedorov [1]. They define non-constructivity as follows: an object x is *non-constructible* from a finite set E of objects if there exists a set $E' \supset E$ that is closed under the operations of adding points, lines, and circles (contains all objects derivable from E') and contains an everywhere dense set of points.

Proposition 2 (Akopyan–Fedorov). *This negative definition is equivalent to the game-theoretic definition given above.*

Proof. The equivalence is essentially proven as [1, Proposition 15, p. 9], but Akopyan and Fedorov (for some reasons) avoid stating explicitly the game-theoretic definition.

Assume that x is non-constructible from E according to the negative definition. Then Bob can prevent Alice from winning by always choosing points from E' when Alice asks for a point in an open set. Since E' is dense, these points are enough. If Bob follows this strategy, then the current position will always be a subset of E' and therefore will never contain x.

On the other hand, assume that x in not constructible from E in the sense of the positive definition. Consider the following strategy for Alice. She takes some triangle abc and point d inside it and ask Bob to add points A, B, C, D that belong to some small neighborhoods of a, b, c, d respectively. The size of these neighborhoods guarantees that ABC is a triangle and D is a point inside ABC. There are two cases:

- for every choice of Bob Alice has a winning strategy in the remaining game;
- there are some points A, B, C, D such that Alice does not have a winning strategy in the remaining game.

In the first case Alice has a winning strategy in the entire game and x is constructible. In the second case we consider the set E' of all objects derivable from $E \cup \{A, B, C, D\}$. As we have seen in the proof of the previous proposition, this set is dense. Therefore, x is non-constructible in the sense of the negative definition.

The advantage of the game definition is that it can be reasonably extended to the uniform case. For the uniform case the game is no more the game with full information. Alice sees only the names (and types) of geometric objects in E, and assigns names to new objects produced by Bob. One should agree also how Alice can get information about the configuration and show she can specify the connected component when asking Bob for a point in this component. For example, we may assume that Alice has access to the list of all connected components and full topological information about the structure they form, and the names and positions of objects from E appearing in this structure. Then Alice may choose some component and request a point from it. To win, Alice needs to specify the name of the required object x. After we agree on the details of the game, we may define construction algorithms as computable strategies for such a game. (Note that in this version Alice deals only with finite objects.)

6 Formal definitions are important

In fact, the absence of formal definitions and exact statements is more dangerous than one could think. It turned out that some classical and well known arguments contain a serious gap that cannot be filled without changing the argument. This happened with a proof (attributed to Hilbert in [5]) that one cannot find the center of a given circle using only a straightedge. It is reproduced in many popular books (see, e.g., [15,6,10]) and all the arguments (at least in the four sources mentioned above) have the same gap. They all go as follows [10, p. 18]:

Let the construction be performed in a plane P_1 and imaging a transformation or mapping T of the plane P_1 into another plane P_2 such that:
(a) straight lines in P_1 transform into straight lines in P_2 $\langle \dots \rangle$
(b) The circumference C of our circle is transformed into a circumference $T(C)$ for some circle in P_2.

As the steps called for in the construction are being performed in P_1, they are being faithfully copied in P_2. Thus when the construction in P_1 terminates in the centre O of C, the "image" construction *must* terminate in the centre $T(O)$ of the circle $T(C)$.

Therefore if one can exhibit a transformation T satisfying (a) and (b), but such that $T(O)$ *is not* the centre of $T(C)$, then the impossibility of constructing the centre of a circle by ruler alone will be demonstrated.

Such a transformation indeed exists, but the argument in the last paragraph has a gap. If we understand the notion of construction in a non-uniform way and require that the point was among the points constructed, the argument does not work since the center of $T(C)$ could be the image of some other constructed point. If we use some kind of the uniform definition and allow tests, then these tests can give different results in P_1 and P_2 (the projective transformation used to map P_1 into P_2 does not preserve the ordering), so there is no reason to expect that the construction is "faithfully copied". And a uniform definition that does not allow tests and still is reasonable, is hard to imagine (and not given in the book). Note also that some lines that intersect in P_1, can become parallel in P_2.

It is easy to correct the argument and make it work for the definition of constructibility given above (using the fact that there are many projective mappings that preserve the circle), but still one can say without much exaggeration that the first correct proof of this impossibility result appeared only in [1]. One can add also that the stronger result about two circles that was claimed by Cauer [5] and reproduced with the same proof in [15], turned out to be plainly false as shown in [1], and the problems in the proof were noted already by Gram [8]. It is not clear why Gram did not question the validity of the classical proof for one circle, since the argument is the same. Gram did not try to give a rigorous definition of the notion of a geometric construction, speaking instead about constructions in "ordered plane" and referring to Bieberbach's book [3] that also has no formal definitions.

The weak version of Cauer's result saying that for some pairs of circles one cannot construct their centers, can be saved and proven for the definition of constructibility discussed above (see [1] and the popular exposition in [17]).

It would be interesting to reconsider the other results claimed about geometric constructions (for example, in [9,19,20,21]) to see whether the proofs work for some clearly defined notion of a geometric construction. Note that in some cases (e.g., for Tietze's results) some definition of the geometric construction for the uniform case is needed (and the negative definition is not enough).

References

1. Akopyan, A., Fedorov, R., *Two circles and only a straightedge*, `https://arxiv.org/abs/1709.02562` (2017)

2. Baston, V.J., Bostock, F.A., On the impossibility of ruler-only constructions, *Proceedings of the American Mathematical Society*, **110** (4), December 1990

3. Bieberbach, *Theorie der geometrischen Konstruktionen*, Springer Basel AG (1952), `DOI10.1007/978-3-0348-6910-2`

4. Borel, E., Le calcul des intégrales définies, *Journal de Mathématiques pures et appliquées*, ser. 6, **8**(2), 159–210 (1912)

5. Cauer, D., Über die Konstruktion des Mittelpunktes eines Kreises mit dem Lineal allein. *Mathematische Annallen*, **73**. S. 90–94. A correction: **74**, S. 462–464 (1913)

6. Courant, R., Robbins, H., revised by Stewart, I., *What is Mathematics? An elementary approach to ideas and methods.* Oxford University Press (1996)

7. Engeler, E., Remarks on the theory of geometrical constructions, *The Syntax and Semantics of Infinitary Languages*, edited by Jon Barwise, Springer (Lecture Notes in Mathematics, **72**), pp. 64–76 (1968)

8. Gram, C., A remark on the construction of the centre of a circle by means of the ruler, *Math. Scand.*, **4**, 157–160 (1956)

9. Hilbert, D., *The Foundations of Geometry*, authorized translation by E.J. Townsend, Ph.D., University of Illinois (1902). Reprint edition: The Open Court Publishing Company, La Salle, Illinois (1950). Available at `https://archive.org/details/thefoundationsof17384gut`

10. Kac, M., Ulam, S.M., *Mathematic and Logic*, Dover publications, 1992, copyright: Encyclopaedia Britannica (1968)

11. Kijne, D., *Plane construction field theory*, Ph.D. thesis, promotor H. Freudenthal, van Gorcum & Co., N.V., G.A.Hak & Dr. H.J. Prakke (28 May 1956)

12. Kutuzov, B.V., *Studies in mathematics*, vol. IV, Geometry (L.I.Gordon and E.S.Shater, transl.) School Mathematics Study Group, University of Chicago, Chicago (1960)

13. Manin, Yu. (Манин, Ю.И.), On the decidability of geometric construction problems using compass and straightedge [Russian], 205–227 (1963)

14. Martin, G.E., *Geometric Constructions*, Springer (1998)

15. Rademacher, H., Toeplitz O., *Von Zahlen und Figuren. Zweite Auflage. Springer-Verlag, Berlin, Heidelberg*, 1933.

16. Schreiber, P., *Theorie der geometrischen Konstruktionen*, VEB Deutscher Verlag der Wissenschaften, Berlin (1975)

17. Shen, A., *Hilbert's Error?*, `https://arxiv.org/abs/1801.04742` (2018)

18. Tao, T., *A geometric proof of the inpossibility of angle trisection*, `https://terrytao.wordpress.com/2011/08/10/a-geometric-proof-of-the-impossibility-of-angle-trisection-by-straightedge-and-compass/`

19. Tietze, H., Ueber dir Konstruierbarkeit mit Lineal und Zirkel, *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, Abt. IIa, 735–757 (1909). A short resume: `https://www.biodiversitylibrary.org/item/93371`

20. Tietze, H., Ueber die mit Lineal und Zirkel und die mit dem rechten Zeichenwinkel lösbaren Konstruktionaufgaben I, *Mathematische Zeitschrift*, **46**, 190–203 (1940), `http://www.digizeitschriften.de/dms/img/?PID=GDZPPN002379074`

21. Tietze, H., Zur Analyse der Lineal- und Zirkelkonstruktionen. I. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Abteilung der Bayrischen Akademie der Wissenschaften zu München, 1944, Heft III, Sitzungen Oktober–Dezember*, 209–231, München (1947). Available at `http://publikationen.badw.de/003900992.pdf`

22. Wantzel, M.L., Recherches sur les moyens de reconnaître si un probléme de Géométrie peut se résoudre avec la règle et le compas, *Journal de mathématiques pures et appliquées*, 1re série, **2**, 366–372 (1837)

Appendix

Due to the page limit, we omitted more detailed analysis of the approaches to non-determinism and arbitrary points used by different authors. Here are some corresponding quotes with comments.

Bieberbach and Tietze on the non-uniformity

Bieberbach [3, p. 26] describes this problem as follows:

> Noch mag aber ausdrücklich hervorgehoben werden, daß es sich in diesem Paragraphen ebenso wie bei den Poncelet–Steinerschen Konstruktionen stets um ein Konstruieren in der orientierten Ebene handelt. Es soll bei jedem gegebenen und bei jedem konstruierten Punkt feststehen, welches die Vorzeichen seiner Koordinaten sind. Anderenfalls steht nur fest, daß der gesuchte Punkt sich unter den konstruierten befindet. [7]

Bieberbach then adds (on p. 151):

> H.TIETZE hat (l.c.,§4) zu dem Ergebnis dieses Paragraph die folgende Bemerkung gemacht: Bei der Ausführung der Konstruktion erhält man im Schnitt eines Kreises mit einer Geraden oder einem anderen Kreis stets mehrere Punkte. Bei der Fortsetzung der Konstruktion bedarf es dann noch einer Angabe darüber, welcher der beiden beim n-ten Schritt erhaltenen Punkte bei der Fortsetzung der Konstruktion benutzt werden soll. Man kann fragen, bei welchen Konstruktionsaufgaben es stets gleichgültig ist, welchen der beim n-ten Schritt erzeugten Punkte man beim $(n+1)$-ten Konstruktionsschritt verwendet. [8]

Indeed, Tietze studied this question in several papers, starting from 1909 [19,20,21]. He noted the necessity of ordering tests for standard constructions already in [19] amd improves the exposition in [20]. Then [21, pp. 228–229] he considers also the question: which *distances* can be constructed without ordering tests and notes that one can construct a distance that is $\sqrt{3}$ times longer that a given one (just consider the distance between the intersection points of two circles with centers A, B and radius AB) but this is not the case, say, for $\sqrt{3}+1$ instead of $\sqrt{3}$.

---

[7] One should say explicitly that the constructions of this paragraph (as well as the Poncelet–Steiner constructions) are performed in the oriented plane. For each given and constructed point it should be determined what the signs of its coordinates are. Otherwise it is only certain that the point in question is among the constructed points.

[8] Tietze [21, §4] considered the problem mentioned in this section and noted that the intersection of a circle with a line or other circle contains several points, and one should specify which or the points obtained at nth step of the construction should be used. One may ask for which constructions it does not matter which point is used for the next $(n+1)$th step.

Different authors about arbitrary points

Tietze does not consider the problem of arbitrary points; in [21] he writes:

> 18. Was wir nicht in unsere Betrachtungen einbezogen haben, sind
> willkürlich gewählte Elemente, wie sie bei manchen Konstruktionen eine
> Rolle spielen. Diese Rolle ist aber nicht ganz so einfach, als es bisweilen
> dargestellt wird, weil beispielsweise eine willkürlich gewählte Gerade, je
> nach der getroffenen Wahl, mit einer gegebenen oder bereits konstruier-
> ten oder später zu konstruierenden Kreislinie (oder auch mit der Figur
> schon angehörenden oder später zu konstruierenden Geraden) Schnitt-
> punkte haben kann oder nicht. ⟨...⟩ Ob es einmal in einer Fortsetzung
> dieser Note zu einer Besprechung auch dieser Dinge kommen wird, muß
> offen bleiben.[9]

Baston and Bostock [2] do not even attempt to give a formal definition of
a constructible point. However, they acknowledge that the use of "arbitrary"
points in necessary for some constructions, and make some (rather vague, to be
honest) remarks about that (p. 1018):

> *Result* 1. Given three distinct collinear points A , B, and C such that
> B is the midpoint of the segment AC, then it is possible using a ruler
> alone to construct the midpoint M of the segment AB.
> Here we see that $L(\{A, B, C\}) = \{A, B, C\}$ [the left-hand side is the clo-
> sure of points A, B, C] so, although the midpoint M is meant to be con-
> structible from the set $\{A, B, C\}$, it is not actually derivable. Of course
> the constructibility of M depends on the (not unreasonable) ability to
> choose arbitrary[1] points not on the line ABC. Although the possibility
> of sets giving rise to constructible points which are not derivable may
> be somewhat disturbing, it can be of some comfort that such sets are
> probably quite limited. Fairly clear evidence will appear later which es-
> sentially indicates that if a set E contains at least four points, not the
> vertices of a parallelogram and no three of them collinear, then the ability
> to choose arbitrary points (with reasonable properties) will not permit
> the construction with ruler alone of any point which is not l-derivable
> [=derivable using a ruler only]. In the case of ruler-and-compasses or
> compasses-only constructions, we suspect the situation is even simpler.
> Although we will not be presenting any real evidence to support the view,
> we do have a strong feeling that, whenever E contains at least two points,
> there are no constructible points which are not then derivable. Thus our
> overall conclusion is that the distinction between constructibility and

---

[9] What we do not include in our analysis are arbitrary elements that are used in some
constructions. They role is not so simple, however, as it is sometimes thought, since,
for example, an arbitrary line may (depending on the choice made) intersect or may
not intersect some circle that is already constructed or will be constructed later; the
same is true for other figures formed by existing or future lines. ⟨...⟩ We may (or
may not) return to this question in the continuation of this note.

derivability arising from the use of arbitrary points is not very complex. Since, in addition, this aspect of constructibility does not actively cause the inconsistencies in the literature to be described presently, we will not pursue a more detailed analysis in this direction.

The footnote "[1]" says "The interested reader may wish to consult [4,p. 79] where an elementary approach to the use of arbitrary points can be found." The reference points to the book [12]; however, the corresponding explanations (p. 46 of the Russian original version) are also far from being clear:

> One often uses *arbitrary* elements in the geometric constructions. The option of adding arbitrary elements to the current configuration needs to be restricted, and there restrictions are usually formulated as follows:
> RESTRICTION A. *One may consider an arbitrary point of the plane outside the given line as constructed.*
> RESTRICTION B. *One may consider an arbitrary point on a given line that differs from the already constructed points in this line, as constructed.*
> The items A and B are essential for many construction problems. ⟨...⟩
> Of course, when proving the correctness of a construction that *uses arbitrary elements* we should not use any special properties of these elements and should consider them as essentially arbitrary points.[10]

Another popular exposition [10, p. 18] explains the use of arbitrary elements (in the context of finding the center of a given circle by a straightedge only) as follows:

> What is a construction by straightedge alone? It is a *finite* succession of steps, each requiring that either a straight line be drawn, or a point of intersection of two lines or of a line and the given circle be found. A straight line can be drawn through two points chosen more or less arbitrarily. For example, a step may call for choosing two arbitrary points on the circumference of the circle and joining them by a straight line. Or it may be drawn through two points, one or both of which have been

---

[10] Часто пользуются при построениях *произвольными элементами*. Возможность приобщения произвольных точек к числу данных или уже построенных нуждается в особых оговорках, которые обычно формулируются в виде следующих требований:

Требование А. *Может считаться построенной произвольная точка плоскости вне данной прямой.*

Требование В. *Может считаться построенной произвольная точка на данной прямой, не совпадающая ни с одной из уже построенных на прямой точек.*

Требования А и В являются существенными при решении многих задач. ⟨...⟩

При доказательстве правильности того или иного построения *с использованием произвольных элементов* мы не должны, конечно, опираться на какие-либо специальные свойства этих элементов, а должны существенным образом считать эти элементы произвольными.

determined in a previous step in the construction as intersections of lines or a line and the circle. The succession of steps eventually must yield a point that can be proved to be the centre of our circle.

Both quotes speak about *proofs* that the constructed point has the desired properties (though do not specify what kind of proofs they have in mind). This is an old theme that goes back to Hilbert. In his classical book on the foundations of geometry [9] there is a chapter "Geometrical constructions based upon the axioms I–V" that speaks, e.g., about "those problems in geometrical constructions, which may be solved by the assistance of only the axioms I–V", but there are no exact definitions of what does it mean. This there was developed by Schreiber [16], but his approach is quite unclear.

Probably the most detailed exposition of the role of arbitrary points is given by Manin in [13] (an encyclopedia of elementary mathematics addressed to advanced high school students and teachers):

It is convenient to describe the construction process inductively. We start with a finite set of points of the plane and want to obtain another finite set of points. The construction process adds some new points to the existing once, and then selects the answer to our problem among the points constructed. The second (selection) part depends on the nature of the problem; now we are interested in the process of a d d i n g new points.

By a *construction step* we mean the process of adding *one* new point. To find this new point, we perform some operations. *B y d e f i n i t i o n*, a compass and straightedge construction may use only the following operations (note that operations 1 and 2 can be used several times while operation 3 is used once during the construction step).

1. *Drawing a line through two points from a current set of points.* (By the current set of points we mean the result of the preceding construction step or the initial set for the first step.)

2. *Drawing a circle whose center is a point from the current set that contains some other point from the current set.* [Manin does not allow to draw a circle that has center A and radius BC where A, B, C are already constructed, but this does not matter much.] ⟨. . .⟩

3. *Choosing one intersection points of the lines and circles constructed and adding this point to the current set.*

Instead of operation 3 where some s p e c i f i c point is added, we sometimes need to use the following operation

3a. *Choosing an "arbitrary" point and adding it to the current set.*

One should specify what do we mean by "arbitrary" point. By definition, it means that the point can be "arbitrarily" chosen either on some line segment, or arc of a circle, or in some part of the plane bounded by segments or arcs (this part can also be unbounded). When doing this, we may use only the lines and circles constructed at the current construction step, and the endpoints of the segments and arc should belong to the current set. We may assume, of course, that the "arbitrary point" is not

an endpoint of a segment/arc or belongs to the interior of the part of the plane where it is chosen.

A compass and straightedge construction is a finite sequence of steps described above. ⟨...⟩

Summarizing, let us assume that a finite set of points of the plane is given. We say that a point *A is constructible* (by compass and straightedge), if there exists a construction such that the resulting set of points contains *A* for all possible intermediate "arbitrary" choices. A construction problem is *solvable* if it requires to find a set of points where each points is constructible by compass and straightedge from given points.[11]

---

[11] Удобно описать процесс построения индуктивно. Мы начинаем с конечного числа точек на плоскости и хотим получить конечное число точек на плоскости; процесс построения состоит в том, что к уже имеющейся системе точек мы добавляем по известным правилам ещё некоторые, а затем отбираем среди всех получившихся точек те, которые доставляют решение нашей задачи. Вторая часть, конечно, определяется спецификой задачи; нас интересует сейчас, по каким правилам д о б а в л я ю т с я точки.

Назовём *шагом построения* добавление *одной* новой точки к уже имеющимся. Отыскание этой новой точки производится в результате проведения некоторых операций; п о о п р е д е л е н и ю, в построении циркулем и линейной шаг может состоять только из следующих операций (причём операции 1 и 2 могут применяться несколько раз, а операция 3 — один раз).

1. *Проведение прямой через пару точек имеющейся совокупности.* (Эта совокупность является результатом предыдущего шага или представляет собой первоначально заданную систему точек.)

2. *Проведение окружности с центром в одной из точек имеющейся совокупности и проходящей через некоторую другую точку этой совокупности.* ⟨...⟩

3. *Выбор одной точки пересечения построенных прямых и окружностей между собой и добавление этой точки к имеющейся совокупности.*

Вместо операции 3, которая представляет собой выбор о п р е д е л ё н н о й т о ч к и, иногда оказывается необходимым пользоваться операцией

3a. *Выбор «произвольной» точки и добавление её к имеющейся совокупности.*

Следует уточнить здесь употребление слова «произвольной». По определению это означает, что точку можно «произвольно» выбирать либо на некотором отрезке прямой, либо на дуге окружности, либо же в части плоскости, ограниченной отрезками или дугами, и, возможно, уходящей в бесконечность. При этом все фигурирующие прямые и окружности должны быть построены на данном шаге, а концы отрезков и дуг должны быть точками имеющейся совокупности. Можно считать, конечно, что «произвольная» точка не должна лежать на концах отрезков и дуг или на границе упомянутой части плоскости.

Построением с помощью циркуля и линейки называется последовательность, состоящая из конечного числа описанных шагов. ⟨...⟩

Итак, пусть задана некоторая конечная совокупность точек на плоскости; мы считаем, что точку *A можно построить* (с помощью циркуля и линейки), если существует такое построение, что (независимо от промежуточных «произвольных» выборов точек!) система точек, полученная в результате этого построения, содержит точку *A*. Задачу на построение мы считаем *разрешимой*, если совокуп-

Note the, being literally understood, this definition makes no sense. Indeed, a "construction" that should exist is a finite sequence of steps as described, so it (if understood in a natural sense) determines the "arbitrary" points that were added during operations of type 3a. So we cannot add the quantifier "for all possible intermediate choices".

ность точек, которые следует найти для решения этой задачи, состоит только из таких точек, которые можно построить с помощью циркуля и линейки.