Topology and its Applications 16 (1983) 19–31
North-Holland

CLASSIFICATION OF KNOT PROJECTIONS

C.H. DOWKER and Morwen B. THISTLETHWAITE
Birkbeck College and Polytechnic of the South Bank, London, England

Received 7 May 1982

The first step in tabulating the non-composite knots with n crossings is the tabulation of the non-singular plane projections of such knots, where two (piecewise linear) projections are regarded as equivalent, or in the same class, if they agree up to homeomorphism of the extended plane, i.e. two-sphere. This first step is here reduced to a simple algorithm suitable for computer use.

AMS Subj. Class.: 57M25

knot	classification of knots
algorithm	tabulation of knots

The nineteenth century knot tables were based on a prior classification of knot projections. The knot projections with 8 to 11 crossings were found, mainly by Kirkman (see e.g. [2]), using geometrical methods. But Tait [3, Part 1] tabulated the knots with up to 7 crossings using mainly combinatorial methods. We have modified Tait's notation and methods to make them convenient for computer programming. This paper contains the theoretical justification of the methods.

Although none of the proofs needs diagrams, the reader will find the paper much easier to follow if he makes sketches.

Let L be a regular projection [1] of a knot k. Each crossing in L is the image of two points, the upper crossing and lower crossing, in K. If there are n crossings in L, and hence $2n$ crossing points in K, we number the crossing points $1, 2, \ldots, 2n$ consecutively, starting from any one of them, and going in a chosen direction along K.

There is an involutory function a from $\{1, 2, \ldots, 2n\}$ to itself, i.e. $a(a(i)) = i$, for which $a(i)$ is the other crossing point of K having the same projection as i. We shall usually write a_i for $a(i)$. The involution a is parity reversing: if i is odd then $a(i)$ is even, and vice versa. We write S for the sequence $a_1, a_2, \ldots, a_{2n}$. The subsequence $a_1, a_3, \ldots, a_{2n-1}$ determines S and a.

There are $4n$ ways of numbering the crossing points of K, corresponding to $2n$ possible starting points and 2 possible directions, and the corresponding sequences need not be all the same. The *standard* sequence corresponding to the knot

S is realizable. Hence, by Theorem 2, S is realizable if and only if (S, f) satisfies Rule 2.

Here we use a slightly different form of Rule 2, justified by the symmetries of $\phi_i(s)\phi_i(a_s)$. When $s<i<a_s<a_i$, let $t=a_i$. Then $\phi_i(s)\phi_i(a_s)=\phi_t(s)\phi_t(a_s)=-\phi_s(t)\phi_s(a_t)$ and $f(i)=-f(t)$, so

$$\phi_i(s)\phi_i(a_s)f(i)f(s)=\phi_s(t)\phi_s(a_t)f(s)f(t).$$

When $s<i<a_i<a_s$, let $t=a_s$. Then

$$\phi_i(s)\phi_i(a_s)=\phi_i(t)\phi_i(a_t)=\phi_t(i)\phi_t(a_i).$$

When $s<a_s<i<a_i$, let $t=a_i$. Then $\phi_i(s)\phi_i(a_s)=\phi_s(i)\phi_s(a_i)=\phi_s(t)\phi_s(a_t)$. In all cases, $i=i_h$ for some h. This completes the proof.

Note that the definition of the canonical orientation and the checking of Rule 2 involve the same function $\phi_i(s)\phi_i(a_s)$. They can conveniently be combined in a single operation.

This completes the theoretical justification of an adequate method of tabulating knot projections. The following remarks leading to Corollary 3 are merely to make the computing somewhat more efficient.

The sequence S may contain a *twist*, that is, a maximal pair of intervals $[r, s]$ and $[a_r, a_s]$ or $[a_s, a_r]$, where $r \neq s$, which are mapped montonically onto each other by the involution a.

Consider first a twist not containing 1 and a_1. There is a first i_h such that $[i, a_i]$, where $i=i_h$, contains part but not all of the twist. Then $[i, a_i]$ contains, say, $[r, s]$, and $[a_i, i]$ contains $[a_r, a_s]$ or $[a_s, a_r]$. Then ϕ_i determines $f(a_q)=\phi_i(a_q)\phi_i(q)f(i)$ for $r \leqslant q \leqslant s$ and, since $\phi_i(q)=\phi_i(r)$ and $\phi_i(a_q)=(-1)^{q-r}\phi_i(a_r)$, therefore $f(a_q)=(-1)^{q-r}f(a_r)$ and $f(q)=(-1)^{q-r}f(r)$.

If 1 is on the twist with, say, $1 \in [r, s]$, and if the mapping a is increasing from $[r, s]$ onto $[a_r, a_s]$, then ϕ_1 determines $f(t)=(-1)^{t-1}$ for $r \leqslant t \leqslant s$. In case a is increasing from $[1, a_1]$ onto $[a_1, s]$, then the whole knot is a twist with $a_1=n+1$, and hence with n odd, and $f(t)=(-1)^{t-1}$ for all t.

If $1 \in [r, s]$ and a is decreasing from $[r, s]$ onto $[a_s, a_r]$ then the orientation at i_2 is determined by ϕ_1 and the orientation on the twist, except at 1 and a_1, is determined by ϕ_i when $i=i_2$. Indeed $i=i_2 \in [1, a_1]$, $a_i \notin [1, a_1]$ and $f(a_i)=\phi_1(a_i)\phi_1(i)f(1)=\phi_1(i)\phi_1(a_i)$. Hence

$$f(i)=-\phi_1(i)\phi_1(a_i)=\phi_i(1)\phi_i(a_1) \quad \text{and} \quad \phi_i(1)\phi_i(a_1)f(i)=1=f(1).$$

Thus the orientation determined by ϕ_i on the rest of the twist agrees at 1 and a_1 with that already determined. Hence $f(t)=(-1)^{t-1}$ for $t \in [r, s]$. Thus in all the cases, f is alternating on the twist.

Now suppose a is increasing from $[r, s]$ onto $[a_r, a_s]$. Let $p \in [r, s]$. Let $t \notin [p, a_p]$ and $a_t \in [p, a_p]$, with $t \notin [r, p-1]$. Then $\phi_p(t)=(-1)^{p-r}\phi_r(t)$, $\phi_p(a_t)=\phi_r(a_t)$, $f(p)=(-1)^{p-r}f(r)$, and hence $\phi_p(t)\phi_p(a_t)f(p)=\phi_r(t)\phi_r(a_t)f(r)$. If $t \in [r, p-1]$, then $\phi_p(t)=$

$(-1)^{p-t}$, $\phi_p(a_t)=1$ and hence $\phi_p(t)\phi_p(a_t)f(p)=f(t)$. If $t\notin[p,a_p]$ and $a_t\notin[p,a_p]$, then $\phi_p(t)\phi_p(a_t)=\phi_r(t)\phi_r(a_t)$. Thus, for defining $f(t)$ or for checking Rule 2, ϕ_p gives nothing not already given by ϕ_r.

If a is decreasing from $[r,s]$ onto $[a_s,a_r]$, let $p\in[r,s]$. Let $t\notin[p,a_p]$ and $a_t\in[p,a_p]$. Then $\phi_p(t)=\phi_r(t)$, $\phi_p(a_t)=(-1)^{p-r}\phi_r(a_t)$ and $\phi_p(t)\phi_p(a_t)f(p)=\phi_r(t)\phi_r(a_t)f(r)$. If $t\notin[p,a_p]$ and $a_t\notin[p,a_p]$ and t is not on the twist, $\phi_p(t)\phi_p(a_t)=\phi_r(t)\phi_r(a_t)$. If $r\leqslant t<p$, $\phi_p(t)\phi_p(a_t)=1$. Again ϕ_p gives nothing new.

In all cases one needs to check only one ϕ_r for each twist. If 1 is on the twist, ϕ_1 will do. This justifies changing the definition of i_h as follows: Let $A_1=[1,a_1]$, $B_1=\emptyset$, $i_1=1$. Inductively, let

$$A_h=A_{h-1}\cup\{s: i_{h-1}*s\},$$

$$B_h=B_{h-1}\cup\{i_{h-1},a(i_{h-1})\}$$
$$\cup\{t,a_t: t\notin[i,a_i],\ a_t\in[i,a_i],\ a_{t-1}=a_t\pm 1\ (\mathrm{mod}\ 2n),\ i=i_{h-1}\},$$

$$i_h=\text{least member of } A_h\backslash B_h \text{ if } A_h\backslash B_h\neq\emptyset.$$

This excludes the possibility, for $i=i_h$, that $a_{i-1}=a_i\pm 1$ (mod $2n$). It ensures that only one ϕ_i is checked for each twist, except that it leaves one unnecessary ϕ_i in certain cases of twists containing 1. We still have $\bigcup_h A_h=\{1,\ldots,2n\}$.

The definition of canonical orientation remains formally the same, but it now depends on the new definition of i_h. But, since the new definition merely excludes certain unnecessary ϕ_i when i is on a twist, the orientation defined is the same as before.

Also Rule 2 remains valid when these unnecessary ϕ_i are excluded. Hence we have the following corollary:

Corollary 3. *Theorem* 3 *remains true with the changed meaning of* i_h.

We have used the canonical orientation only for checking whether S is realizable. It will also be needed for finding a presentation of the knot group, and hence for finding the knot invariants. In fact one can find the group and its invariants for an arbitrary orientation of S, whether S is realizable or not, and for an arbitrary choice of overcrossings [1, p. 86]. This enables one to compare knot groups with other similarly presented groups.

The canonical orientation f for non-realizable S may also be of interest. For example, the sequence S whose standard subsequence is 4 8 2 10 6 is not realizable because (S,f) does not satisfy (i) of Rule 2, where f is the canonical orientation: $f(i)=(-1)^{i-1}$. Taking the odd numbers as overcrossings, (S,f) has the same Alexander polynomial as the reef knot and certain non-composite knots. But the group of this 'knot' is distinguished by its invariants from the group of any knot with up to 12 crossings. This (S,f) can be realized on the torus, and the realization is the projection of an alternating knot in $T^2\times\mathbb{R}^1$.

However, if the aim is to classify knots, we need the canonical orientation only for knot projections, and we should stop the computation as soon as we find that (S, f) does not satisfy Rule 2, as in the following algorithm.

Algorithm

Given S, take $A = \{1, a_1\}$, $f(1) = 1$, $f(a_1) = -1$
If $A \neq \emptyset$, select i = least member of A
Take $\phi_i(i) = 1$
If $\phi_i(x-1)$ has been found, where $x-1$ is taken mod $2n$,
  If $a_x \notin [i, a_i]$, $\phi_i(x) = \phi_i(x-1)$
  If $a_x \in [i, a_i]$, $\phi_i(x) = -\phi_i(x-1)$
End with $\phi_i(i-1)$
Take $D_i = \{1, 2, \ldots, 2n\} \backslash [i, a_i]$
  If $D_i \neq \emptyset$, select x = least member of D_i
  If $x < i$
    If $a_x \notin [i, a_i]$
      If $\phi_i(x)\phi_i(a_x) = 1$, remove x and a_x from D_i
      If $\phi_i(x)\phi_i(a_x) = -1$, reject S, proceed to new S
    If $a_x \in [i, a_i]$
      If $f(x)$ is already defined
        If $\phi_i(x)\phi_i(a_x)f(i) = f(x)$, remove x from D_i
        If $\phi_i(x)\phi_i(a_x)f(i) = -f(x)$, reject S, proceed to new S
      If $f(x)$ not already defined, let $f(x) = \phi_i(x)\phi_i(a_x)f(i)$, $f(a_x) = -f(x)$
        If $a_{x-1} = a_x \pm 1$ ($a_x + 1$ is mod $2n$), remove x from D_i
        If $a_{x-1} \neq a_x \pm 1$, add x and a_x to A, remove x from D_i
  If $x > a_i$
    If $a_x \notin [i, a_i]$, remove x and a_x from D_i
    If $a_x \in [i, a_i]$
      If $f(x)$ is already defined, remove x from D_i
      If $f(x)$ not already defined, let $f(x) = \phi_i(x)\phi_i(a_x)f(i)$, $f(a_x) = -f(x)$
        If $a_{x-1} = a_x \pm 1$ ($a_x + 1$ is mod $2n$), remove x from D_i
        If $a_{x-1} \neq a_x \pm 1$, add x and a_x to A, remove x from D_i
  If $D_i \neq \emptyset$, select x = least member of D_i, etc.
  If $D_i = \emptyset$, remove i and a_i from A
If $A \neq \emptyset$, select i = least member of A, etc.
If $A = \emptyset$, record that (S, f) is realizable, proceed to new S.

Thus tabulating knot projections with n crossings, and their orientations, is algorithmic. We also use algorithms for listing the knots with n crossings, but the resulting list may have duplicates. Our methods of eliminating the duplicates, and projection is the one for which the subsequence $a_1, a_3, \ldots, a_{2n-1}$ is minimal in lexicographic order. The standard sequences corresponding to the knot projections with up to 5 crossings are determined by the subsequences:

$$4\,6\,2, \quad 4\,6\,8\,2, \quad 4\,8\,10\,2\,6, \quad 6\,8\,10\,2\,4.$$

Now let S be any sequence $a_1, a_2, \ldots, a_{2n}$, such that $a : i \to a_i$ is a parity reversing involution of $\{1, 2, \ldots, 2n\}$. Our main aim in this paper is to find out which sequences of this kind correspond to knot projections. This will enable us to use a computer to tabulate knot projections, and eventually to tabulate knots.

Here we should state that there are several solutions in the literature to Gauss's "crossing sequence problem", which is equivalent to our problem of deciding which sequences S correspond to knot projections in the plane (see [4, 5, 6]). However, we develop an alternative approach which is very elementary in nature, and which leads to a simple and efficient algorithm suitable for the computer enumeration of large numbers of knot projections.

Let $[i, j]$ denote the interval $\{i, i+1, \ldots, j\}$ of integers mod $2n$. Let the sequence S correspond to a projection of a knot K. If an interval $[i, i+1] \bmod 2n$ with only two elements were mapped onto itself by a, i.e. if $a_i = i+1 \pmod{2n}$, the projection would be immediately reducible to one with $n-1$ crossings. Thus adjacent numbers $i, i+1 \pmod{2n}$ cannot be interchanged by the involution a. It follows that n cannot be 1 or 2. If K can be projected with no crossings, it is regarded as unknotted. Thus S cannot be the trivial sequence for which $n = 0$.

If there is a proper subinterval $[i, j] \bmod 2n$ of $\{1, 2, \ldots, 2n\}$ which is mapped onto itself by a, the complementary interval $[j+1, i-1] \bmod 2n$ is also mapped onto itself. Each of these intervals has an even number of elements. As we have seen, if S comes from a projection of K, neither interval has two elements. Moreover, if an interval with four elements were mapped onto itself, two adjacent numbers mod $2n$ would be interchanged, which is not possible. Hence K is the composite of two knots each with at least three crossings. When tabulating knots one usually omits composite knots. This justifies imposing the following restrictions on the sequences to be used when tabulating knots or knot projections.

Rule 1. *The sequence $a_1, \ldots, a_{2n}$ is to satisfy*:

(i) $n \geq 3$;

(ii) *no proper subinterval* $[i, j] \bmod 2n$ *of* $\{1, 2, \ldots, 2n\}$ *is mapped onto itself by the involution* $a : k \to a_k$.

There are $2n(2n-1)$ proper subintervals mod $2n$ of $\{1, 2, \ldots, 2n\}$ with even numbers of elements, but each is the complement of another. Each is mapped onto itself if and only if it is mapped into itself. To check whether a given sequence S satisfies Rule 1, it is sufficient to check $n(n-1)$ intervals mod $2n$, each having at most n elements, or to check the $n(n-1)$ proper subintervals of $[1, 2n]$ not containing $2n$.

From now on, we consider only sequences satisfying Rule 1.

When tabulating knot projections, one need use only standard sequences, with the subsequence $a_1, a_3, \ldots, a_{2n-1}$ minimal in lexicographic order when the transformations $y = b + x$ and $y = b - x$, $b = 1, 2, \ldots, 2n$, are applied to i and a_i. However in this paper it will not be assumed that the sequences are standard.

Now let $[0, 2n]$ denote the closed interval in $\mathbb{R}$. A *realization* of the sequence S is a piecewise linear mapping $\rho : [0, 2n] \to \mathbb{R}^2$ such that

(i) $\rho(0) = \rho(2n)$;

(ii) $\rho(i) = \rho(a_i)$ for $i = 1, 2, \ldots, 2n$;

(iii) ρ maps $[0, 2n] \backslash \{0, 1, \ldots, 2n\}$ homeomorphically;

(iv) the arc $\rho([i - \frac{1}{2}, i + \frac{1}{2}])$ crosses the arc $\rho([a_i - \frac{1}{2}, a_i + \frac{1}{2}])$ at $\rho(i)$, where $i - \frac{1}{2}$, $i + \frac{1}{2}$, etc. are mod $2n$.

For each such realization, there exists an alternating knot K which can be parametrized by $[0, 2n]$ so that ρ is the projection of K to $\mathbb{R}^2$. If there exists a realization of S, we say that S is *realizable*.

Let G be the 4-valent graph obtained from $[0, 2n]$ by identifying 0 with $2n$ and then identifying i with a_i for $i = 1, 2, \ldots, 2n$. It follows from Rule 1 that

(α) each edge of G joins two different vertices;

(β) if any two of its edges are cut, G remains connected;

(γ) if any vertex is removed, G remains connected.

A realization of S may be regarded as a piecewise linear embedding of G in $\mathbb{R}^2$ which preserves or reverses the cyclic order of the edges $[i-1, i]$, $[a_i - 1, a_i]$, $[i, i+1]$, $[a_i, a_i + 1]$ at each vertex i. Or more simply, a realization of S is a piecewise linear embedding $\tau : G \to \mathbb{R}^2$ which preserves the unoriented cyclic order at each vertex.

Let S^2 be the 2-sphere consisting of $\mathbb{R}^2$ with a point at ∞. Let G be any finite graph. Two piecewise linear embeddings $\tau_1, \tau_2 : G \to \mathbb{R}^2$ are called *equivalent* if there is a homemorphism $h : S^2 \to S^2$ such that $h \circ \tau_1 = \tau_2$. Hence we say that two realizations $\rho_1, \rho_2 : [0, 2n] \to \mathbb{R}^2$ of S are *equivalent* if there is a homeomorphism $h : S^2 \to S^2$ such that $h \circ \rho_1 = \rho_2$.

If K_1 is a knot parametrized by $[0, 2n]$ so that ρ_1 is its projection to $\mathbb{R}^2$, and if ρ_2 is equivalent to ρ_1, then there is a parametrized knot K_2, of the same type as K_1, whose projection to $\mathbb{R}^2$ is ρ_2.

Theorem 1. *If S is realizable, any two realizations of S are equivalent.*

This is a consequence of the following lemma.

Lemma 1. *Let G be a finite graph such that (α) each edge of G joins two different vertices; (β) if any two of its edges are cut, G remains connected; (γ) if any vertex is removed, G remains connected. Let there be given an unoriented cyclic order of edges at each vertex of G. If there exists a piecewise linear embedding of G in $\mathbb{R}^2$ which preserves the unoriented cyclic order at each vertex, then any two such embeddings are equivalent.*

Proof. Ignoring trivial cases, we assume throughout that G has at least two edges. Then G is connected and remains connected when one edge is cut.

The proof is inductive. We construct a strictly increasing sequence, necessarily ending with G, of subgraphs Γ satisfying the following conditions:

(1) Γ has a piecewise linear embedding in $\mathbb{R}^2$, preserving the given unoriented cyclic order at each vertex, which is unique up to homeomorphism of S^2.

(2) The embedded Γ, as a subset of S^2, is the 1-dimensional skeleton of a non-degenerate cell decomposition of S^2.

Here non-degenerate means that the intersection of any two(closed) 2-cells is empty or a 0-cell or a 1-cell, and that each 1-cell is the intersection of two 2-cells. It follows that each 0-cell is on at least three 2-cells and on at least three 1-cells, and must be a vertex of the graph Γ. It is not excluded that two 1-cells may have the same boundary 0-sphere.

To start the induction we first choose a vertex v of G. Since G is connected and has edges, v is one end of some edge. Let u be the other end. Cutting the edge between v and u leaves G connected. Choose a minimal path (least number of edges) from v to u in the cut graph. It is an arc and, with the original edge from v to u, forms a circle (simple closed curve). Cutting the two edges at v on this circle leaves G connected. Join v to the rest of the circle by a minimal path in the cut graph. It is an arc meeting the circle only at v and at its other end w, which is a vertex. The subgraph Γ_0 of G consisting of this arc and the circle is homeomorphic to a graph H consisting of two vertices v, w and three edges, each joining v to w. This graph H, and hence also Γ_0, has a unique piecewise linear embedding in $\mathbb{R}^2$, up to homoemorphism of S^2. It is the 1-skeleton of a cell decomposition of S^2 with three 2-cells, three 1-cells and two 0-cells. Since each vertex of Γ_0 has valency at most 3, the preservation of unoriented cyclic order is trivial. Thus Γ_0 satisfies (1) and (2).

Now take any proper subgraph Γ of G which satisfies (1) and (2).

Case 1. There is a vertex v of Γ which is embedded as a 0-cell of the cell decomposition, and which is an end of an edge of G not in Γ.

Let the other end of this edge be u. If u is not in Γ, there is a minimal path, hence an arc, in the connected set $G\backslash\{v\}$ from u to $\Gamma\backslash\{v\}$, ending, say, at w. If $u \in \Gamma$, take $w = u$. In either case we have an arc from v to w through u, with $w \neq v$ and with only the ends of the arc vw in Γ. Adding the edges and vertices on this arc to Γ, we obtain Γ_1.

G has a piecewise linear embedding in $\mathbb{R}^2$ preserving unoriented cyclic order at its vertices. Hence there is an embedding of Γ which extends to a piecewise linear embedding of Γ_1 preserving unoriented cyclic order at its vertices.

The arc from v to w, apart from its endpoints, must be embedded in some open 2-cell C of the cell decompsition corresponding to Γ. The 0-cell v is on the boundary of at least three 1-cells separating at least three 2-cells. Thus v is on at least three edges of Γ. Hence the unoriented cyclic order of the edges of Γ_1 at v determines the 2-cell C in which vu and vw lie.

Let τ and τ' be two piecewise linear embeddings of Γ_1 in $\mathbb{R}^2$ preserving the unoriented cyclic order of edges at each vertex. Since Γ satisfies (1) there is a homeomorphism $h:S^2 \to S^2$ such that $h \circ \tau$ agrees with τ' on Γ. If C is the 2-cell which must contain $\tau(vw)$, then $h(C)$ is the 2-cell which must contain $\tau'(vw)$. There is a homeomorphism $h_1:S^2 \to S^2$ which agrees with h on the complement of the interior of C and agrees with $\tau' \circ \tau^{-1}$ on $\tau(vw)$ and hence on all of $\tau(\Gamma_1)$. Then $h_1 \circ \tau = \tau' : \Gamma_1 \to S^2$, so Γ_1 satisfies (1).

If w is a 0-cell of the embedded Γ, we change the cell decomposition by introducing a new 1-cell vw, which divides the 2-cell C into two 2-cells. If w is not a 0-cell, it lies inside some 1-cell. We divide the 1-cell, at a new 0-cell w, into two 1-cells. Then the new 1-cell vw divides C into two 2-cells. In either case we have a new non-degenerate cell decomposition whose 1-skeleton is Γ_1. Thus Γ_1 satisfies (2).

Case 2. The assumption of Case 1 does not hold but there is a 1-cell containing more than one edge of Γ.

We cut the two end edges of this 1-cell. Since G with these cuts is still connected, we take a minimal arc in the cut graph G from the interior of the 1-cell to the rest of Γ, say from a vertex v inside the 1-cell to a vertex w of Γ not inside the 1-cell. This arc has only its end points in Γ. Since Γ is not in Case 1, w is not a 0-cell but, like v, it lies inside a 1-cell. Hence w is not even an end point of the 1-cell through v. Adding the arc vw to Γ, we obtain Γ_1.

The embedding of G gives us an embedding of Γ which extends to a piecewise linear embedding of Γ_1 preserving unoriented cyclic order at its vertices.

The arc vw, apart from its end points, must be embedded in some open 2-cell C. The 1-cell containing v is the intersection of two 2-cells, one of which must be C. Since w is in C but not in this 1-cell, w is not in the other 2-cell. The arc vw can be embedded only in the cell C. As in Case 1, it follows that Γ_1 satisfies (1).

The 1-cell containing v is divided into two 1-cells by a new 0-cell at v. The 1-cell containing w is divided into two 1-cells by a new 0-cell at w. Then C is divided into two 2-cells by the new 1-cell vw. We have a new non-degenerate cell decomposition of S^2 whose 1-skeleton is Γ_1. Thus Γ_1 satisfies (2).

There are no other cases because, since Γ is a proper subgraph of G and G is connected, there is an edge of G which is not in Γ but has an end point v in Γ. If v is a 0-cell, we are in Case 1. Otherwise, v is inside a 1-cell which must have more than one edge of Γ, so we are in Case 1 or in Case 2.

Thus there exists a subgraph Γ_0 of G satisfying (1) and (2). And for each proper subgraph Γ of G satisfying these conditions, there is a larger subgraph Γ_1 satisfying the conditions. Hence G satisfies (1) and (2). In particular, G satisfies (1), which proves the lemma, and hence also the theorem.

Corollary 1.1. *If S is realizable and G is the 4-valent graph corresponding to S, there is a piecewise linear embedding of G in the oriented plane $\mathbb{R}^2$ which preserves the cyclic order of the edges at vertex* 1, *and preserves the unoriented cyclic order at* *the other vertices. This embedding is unique up to orientation preserving homeomorphism of S^2.*

Note that if two directed arcs cross each other in the oriented plane, the first crosses the second from right to left if and only if the second crosses the first from left to right. Let $\rho:[0, 2n]\to\mathbb{R}^2$ be a realization of S in the oriented plane. We identify 0 with $2n$ and use numbers mod $2n$. Let $f(i)=1$ if the arc $\rho([a_i-1, a_i+1])$ crosses the arc $\rho([i-1, i+1])$ from right to left, $f(i)=-1$ otherwise. Then $f(i)=-f(a_i)$. We can choose the orientation of $\mathbb{R}^2$, or the realization ρ, so that $f(1)=1$.

For any S, realizable or not, we define an *orientation* to be a function $f:\{1, 2, \ldots, 2n\}\to\{-1, 1\}$ such that $(i) f(1)=1$; $(ii) f(i)=-f(a_i)$ for $i=1, 2, \ldots, 2n$.

If f is an orientation of S, a *realization* of (S, f) is a realization ρ of S in the oriented plane $\mathbb{R}^2$ such that $\rho([a_i-1, a_i+1])$ crosses $\rho([i-1, i+1])$ from right to left if and only if $f(i)=1$. (S, f) is called *realizable* if there exists a realization of (S, f).

The following is a restatement of Corollary 1.1.

Corollary 1.2. *If S is realizable, there is a unique orientation f such that (S, f) is realizable, and the realization of (S, f) is unique up to orientation preserving homeomorphism of S^2.*

For any S, we define a sequence of functions $\phi_i:\{1, 2, \ldots, 2n\}\to\{-1, 1\}$ as follows:
(i) $\phi_i(i)=1$;

(ii) $$\phi_i(r)=\begin{cases}-\phi_i(r-1) & \text{if } a_r\in\{i, \ldots, a_i\},\\ \phi_i(r-1) & \text{otherwise.}\end{cases}$$

This is consistent because $\phi_i(r)$ changes sign an even number of times before returning to $\phi_i(i)=1$. We have $\phi_i(a_i)=-1$ and $\phi_i(i-1)=-1$.

The *loops* of the realization ρ are the closed paths $\rho([i, a_i])$. We use the following convention. The component of the complement of the loop $\rho([i, a_i])$ which contains the edges $\rho([i-1, i])$ and $\rho([a_i, a_i+1])$, except their end points, is said to be *outside* the loop, and the components reached from this one by crossing the loop an even number of times are also *outside*. The other components are *inside*.

By abuse of language we write $[i, a_i]$ for the loop $\rho([i, a_i])$. Then $\phi_i(a_i)=-1$ where the path $\rho([0, 2n])$ passes outside the loop $[i, a_i]$. When this path crosses from the outside to the inside, or vice versa, $\phi_i(r)$ changes sign. Thus if $r\in\{a_i, a_i+1, \ldots, i-1\}$, $\phi_i(r)=-1$ if the path at r or just beyond is outside, $\phi_i(r)=1$ if it is inside.

Note that if $\rho(r)$ is not on the loop then, since $\rho(r)$ and $\rho(a_r)$ are the same point, they are both inside or both outside. Hence $\phi_i(r)\phi_i(a_r)=1$.

If $f(i)=1$, the arc through a_i crosses the arc through i from right to left, and the component to the right of $\rho([i, i+1])$ is inside. Also, since $\phi_i(i)=1$, $\phi_i(i)f(i)=1$. If $\phi_i(i)f(i)=-1$, the component to the left of $\rho([i, i+1])$ is inside. Where the loop crosses itself, $\phi_i(r)f(i)$ changes sign and the inside changes from left to right, or vice versa. Thus if $r\in\{i, \ldots, a_i-1\}$, $\phi_i(r)f(i)=1$ if the inside is to the right at r, or just beyond; $\phi_i(r)f(i)=-1$ if the inside is to the left.

Now let $r \notin [i, a_i]$, $a_r \in [i, a_i]$. If the arc through r crosses the loop from outside to inside and the inside is to the right at r, or if it crosses from inside to outside and the inside is to the left at r, then $\phi_i(r)$ and $\phi_i(a_r)f(i)$ have the same sign and hence $\phi_i(r)\phi_i(a_r)f(i) = 1$. Also since the arc through r crosses from left to right, $f(a_r) = -1$ and $f(r) = 1$. Similarly, if $\phi_i(r)\phi_i(a_r)f(i) = -1$, then $f(r) = -1$. Thus $\phi_i(r)\phi_i(a_r)f(i) = f(r)$.

Thus if (S, f) is realizable, (S, f) must satisfy the following condition:

Rule 2. *For all i and s in $\{1, \ldots, 2n\}$ such that $i < a_i < s$ and $a_s < s$,*
(i) $\phi_i(s)\phi_i(a_s) = 1$ *if* $a_s \notin [i, a_i]$;
(ii) $\phi_i(s)\phi_i(a_s)f(i) = f(s)$ *if* $a_s \in [i, a_i]$.

The conditions $i < a_i < s$ and $a_s < s$ make Rule 2 easier to check. They are justified by Theorem 2 below. They can also be justified by the symmetries of the function $\phi_i(s)\phi_i(a_s)$. Obviously $\phi_i(s)\phi_i(a_s) = \phi_i(a_s)\phi_i(s)$. It can be verified that if i, a_i separate s, a_s in the cyclic order mod $2n$,

$$\phi_i(s)\phi_i(a_s) = \phi_{a(i)}(s)\phi_{a(i)}(a_s) \quad \text{and} \quad \phi_i(s)\phi_i(a_s) = -\phi_s(i)\phi_s(a_i).$$

If the cyclic order is i, a_i, s, a_s, then

$$\phi_i(s)\phi_i(a_s) = -\phi_{a(i)}(s)\phi_{a(i)}(a_s) \quad \text{and} \quad \phi_i(s)\phi_i(a_s) = \phi_s(i)\phi_s(a_i).$$

For example, if the cyclic order is i, a_s, a_i, s, $\phi_i(s)\phi_i(a_s) = (-1)^u$, where u is the number of elements t of $[s+1, a_s]$ for which $a_t \in [i, a_i]$. And $\phi_s(i)\phi_s(a_i) = (-1)^v$, where v is the number of elements t of $[s, a_s]$ for which $a_t \in [i+1, a_i]$. Since $a_s \in [i, a_i]$ but $a_i \notin [s, a_s]$, $v = u + 1$ and hence $\phi_i(s)\phi_i(a_s) = -\phi_s(i)\phi_s(a_i)$.

Theorem 2. *Rule 2 is a necessary and sufficient condition for the realizability of (S, f).*

Proof. Necessity has been shown. It remains to be proved that if (S, f) is not realizable, there exist i and s such that $i < a_i < s$ and $a_s < s$ and either
(i) $a_s \notin [i, a_i]$ and $\phi_i(s)\phi_i(a_s) = -1$, or
(ii) $a_s \in [i, a_i]$ and $\phi_i(s)\phi_i(a_s)f(i) = -f(s)$.

Let G be the 4-valent graph obtained from $[0, 2n]$ by identifying 0 with $2n$ and identifying each i with a_i. Let H_r be the image of the closed interval $[1, r]$ in G. There is some r such that H_r has a piecewise linear embedding in $\mathbb{R}^2$ with the orientations where it crosses itself consistent with f, but H_{r+1} does not have such an embedding. Hence the embedding of H_r cannot be extended by a piecewise linear embedding of $[r, r+1]$ so that the resulting embedding of H_{r+1} is consistent with f. Clearly it would be possible to extend to $[r, r+1]$ if $a_{r+1} \notin H_r$; hence $a_{r+1} \leq r$.

Let α, β be points of the edge $[r, r+1]$ with $r < \alpha < \beta < r+1$. We can extend the embedding of H_r to include a linear embedding of $[r, \alpha]$. We can also embed $[\beta, r+1]$, with $r+1$ at a_{r+1}, but otherwise not meeting H_r, so that the orientation at $r+1$ is consistent with $f(r+1)$. Since $[r, r+1]$ cannot be embedded, the points α and β of $\mathbb{R}^2$ cannot be joined by an arc in $\mathbb{R}^2 \backslash H_r$.

By a suitable bilinear mapping of S^2, regarded as the extended complex plane, we can transform α and β to 0 and ∞ respectively. Let C be a minimal subgraph of H_r, in the sense of least number of edges, which separates 0 from ∞. Then C is a simple closed curve with 0 in the bounded component of its complement. If C is suitably oriented, the winding number of C about 0 is 1. That is, as z goes around C in the positive direction, $\arg z$ increases by 2π.

Let the successive vertices of the directed simple closed curve C be $v_1, v_2, \ldots, v_k$ with edges e_j joining v_j to v_{j+1}, where $v_{k+1} = v_1$. Let $\arg_C z$ be defined unambiguously on C as follows: $0 \leqslant \arg_C v_1 < 2\pi$, and $\arg_C z$ is continuous on C except for a jump of -2π at v_1, with the restriction of $\arg_C z$ to $[v_1, v_2]$ continuous at v_1.

Let $\arg_H z$ be defined on H_r as follows: $0 \leqslant \arg_H z < 2\pi$ at the vertex 1, and $\arg_H z$ varies continuously along the path [1, r]. This defines $\arg_H z$ ambiguously at crossing points of H_r, but unambiguously on the edges.

Let $2\pi d_j$ be the jump of $\arg_H z$ as z passes from e_{j-1} to e_j at v_j. Let $2\pi c_j$ be the jump of $\arg_C z$ at v_j. If we change the function $\arg_H z$ by adding $2\pi a$ to its values on e_j, this will increase d_j by a and decrease d_{j+1} by a. Thus $\sum_1^k d_j$ remains unchanged. Changing $\arg_H z$ on each e_j, so that it agrees with $\arg_C z$ on all the edges of C, leaves $\sum_1^k d_j$ unchanged. Therefore $\sum_1^k d_j = \sum_1^k c_j = -1$. Hence some d_j is odd.

Since $\arg_H z$ is continuous along the path $[1, r]$, and d_j is odd, and hence not 0, v_j is a crossing point of H_r. Let it be the crossing point (i, a_i), where $1 \leqslant i < a_i \leqslant r$. Then $\arg_H z$ is continuous on the closed path $[i, a_i]$ except for a jump of $\pm 2\pi d_j$ at i. Hence the winding number of this closed path about the origin 0 is odd. Hence an arc from 0 to ∞ crosses $[i, a_i]$ an odd number of times.

Returning to the piecewise linear embedding of H_r in $\mathbb{R}^2$, an arc from α to β crosses the loop $[i, a_i]$ an odd number of times. Hence one of α, β is inside the loop, and the other is outside.

Since H_r is embedded with the orientations given by f at its crossing points, and since the path $[1, r]$ includes the loop $[i, a_i]$, the geometric interpretation of ϕ_i deduced above is valid on H_r, and ϕ_i satisfies Rule 2 at the crossing points of H_r.

Now suppose, for example, that α is inside the loop $[i, a_i]$ and β is outside. Then either r is inside the loop, or the path crosses inside the loop at r. In either case, $\phi_i(r) = 1$. Since β is outside, either a_{r+1} is outside or the path $[\beta, r+1]$ is crossing from the outside to the inside at $r+1$. Let $s = r+1$.

First suppose a_s is outside. Then a_s is not in $[i, a_i]$, so $\phi_i(s) = \phi_i(r+1) = \phi_i(r) = 1$. Since $a_s \in H_r$ and a_s is outside, $\phi_i(a_s) = -1$. Thus $\phi_i(s)\phi_i(a_s) = -1$.

Now suppose a_s is on the loop and $[\beta, s]$ is crossing from the outside to the inside. Since a_s is on the loop, $\phi_i(s) = \phi_i(r+1) = -\phi_i(r) = -1$. If $\phi_i(a_s)f(i) = 1$, the inside is on the right at a_s, so $[\beta, s]$ is crossing the path through a_s from left to right, so $f(a_s) = -1$ and $f(s) = 1$. If $\phi_i(a_s)f(i) = -1$, the inside is on the left at a_s, so $[\beta, s]$ is crossing the path through a_s from right to left, so $f(s) = -1$. Thus in each case $\phi_i(s)\phi_i(a_s)f(i) = -f(s)$.

Suppose, if possible, that $r = 2n$. Since $H_r = H_{2n}$ and Rule 2 is satisfied in H_r, if $r+1 = 1 \notin [i, a_i]$, and $a_1 \notin [i, a_i]$, then $\phi_i(1)\phi_i(a_1) = 1$. Also if $1 \notin [i, a_i]$ and $a_1 \in [i, a_i]$, then $\phi_i(1)\phi_i(a_1)f(i)=f(1)$. If $i=1$, then $[2n,1]$ comes from outside to 1 and $\phi_1(r)=\phi_1(2n)=-1$, so both α and β would have to be outside. Hence $r<2n$ and $s\leq 2n$. Thus $1\leq i<a_i<s\leq 2n$ and $1\leq a_s<s\leq 2n$. This completes the proof.

Lemma 3.1. *Let B be a proper subset of $\{1,2,\dots,2n\}$ which is mapped onto itself by the involution a, and let C be the complementary set. Then some pair i, a_i in B separates some pair s, a_s in C in the cyclic order* mod $2n$.

Proof. B is divided by C into intervals $B_1,B_2,\dots,B_k$ which in this order are separated by successive intervals $C_1,C_2,\dots,C_k$ (mod $2n$);

$$\{1,\dots,2n\}=B_1\cup C_1\cup B_2\cup C_2\cup\cdots\cup B_k\cup C_k.$$

By Rule 1, there exists $i\in B_1$ with $a_i\notin B_1$. Choose r minimal positive so that some i in some B_h or C_h has a_i in B_{h+r} or C_{h+r}. We may assume $i\in B_h$, $a_i\in B_{h+r}$. By Rule 1, some $s\in C_h$ has $a_s\notin C_h$. By the minimality of r, $a_s\notin C_{h+1}\cup\cdots\cup C_{h+r-1}$. Hence $a_s\notin[i,a_i]$, so i, a_i separate s, a_s.

Let us write $i*s$ to mean that i, a_i separate s, a_s. We define inductively subsets A_h,B_h of $\{1,\dots,2n\}$ and elements $i_h\in A_h$ as follows, starting from $A_1=\{1,a_1\}$, $B_1=\emptyset$, $i_1=1$. For $h>1$, let $A_h=A_{h-1}\cup\{s: i_{h-1}*s\}$, $B_h=B_{h-1}\cup\{i_{h-1},a(i_{h-1})\}$, and let i_h be the least member of $A_h\backslash B_h$ if this set is not empty.

Lemma 3.2. *The sequence $i_1,i_2,\dots$ ends with i_n.*

Proof. Since, from the definition, all i_h and $a(i_h)$ are distinct, there cannot be more than n values of h. Clearly i_1 exists. By Lemma 3.1, since for $1<h\leq 2n$, B_h is a proper subset of $\{1,\dots,2n\}$, there is some $i\in B_h$ and $s\notin B_h$ such that i, a_i separate s, a_s. If $i<a_i$, $i\in\{i_1,\dots,i_{h-1}\}$. Hence s and a_s are in A_h and hence in $A_h\backslash B_h$. So i_h exists for $h\leq n$.

For each S, whether realizable or not, we define a canonical orientation f inductively as follows. Let $f(1)=1$, $f(a_1)=-1$. For s and a_s in $A_h\backslash A_{h-1}$, we may assume that $s\notin[i,a_i]$, $a_s\in[i,a_i]$, where $i=i_{h-1}$. Then let $f(s)=\phi_i(s)\phi_i(a_s)f(i)$ and $f(a_s)=-f(s)$. This defines f on $\bigcup_1^n A_h\backslash A_{h-1}=A_n=\{1,\dots,2n\}$.

Theorem 3. *S is realizable if and only if (S,f) is realizable, where f is the canonical orientation. Hence S is not realizable if and only if there exist h and s such that, for $i=i_h$, $s<i$ and either $a_s\in[i,a_i]$ and $\phi_i(s)\phi_i(a_s)f(i)=-f(s)$* or *$s<a_s\notin[i,a_i]$ and $\phi_i(s)\phi_i(a_s)=-1$.*

Proof. If S is realiable then, by Corollary 1.2, there is a unique orientation g such that (S,g) is realizable. We have $f(1)=1=g(1)$ and $f(a_1)=-1=g(a_1)$. Assume that $f(j)=g(j)$ for all $j\in A_{h-1}$. For $s\in A_h\backslash A_{h-1}$ with $s\notin[i,a_i]$, where $i=i_{h-1}$, $f(s)=\phi_i(s)\phi_i(a_s)f(i)=\phi_i(s)\phi_i(a_s)g(i)=g(s)$ and $f(a_s)=-f(s)=-g(s)=g(a_s)$. Thus, by induction, $f=g$. Hence (S,f) is realizable. Trivially, if (S,f) is realizable, verifying that they have been eliminated, are partly empirical. The methods used for tabulating knots will be described elsewhere.

A list of knots with up to 12 crossings with sufficient invariants to distinguish them from each other, is in preparation. The knots with 13 crossings are now being computed.

References

[1] R.H. Crowell and R.H. Fox, Introduction to Knot Theory (Ginn and Company, New York, 1963).

[2] T.P. Kirkman, The enumeration, description and construction of knots of fewer than ten crossings, Trans. Roy. Soc. Edinburgh 32 (1885) 281–309.

[3] P.G. Tait, On Knots, Trans. Roy. Soc. Edinburgh 28 (1877) 145–190; also in Scientific Papers, Vol. 1 (Cambridge University Press, 1898).

[4] M. Dehn, Uber Kombinatorische Topologie, Acta Math. 67 (1936) 123–168.

[5] R.C. Read and P. Rosenstiehl, Principal edge tripartition of a graph, Annals of Discrete Math. 3 (1978) 195–226.

[6] R.C. Read and P. Rosenstiehl, On the Gauss crossing problem, Proc. Fifth Hungarian Comb. Colloq., 1977.