Graduate Texts in Mathematics

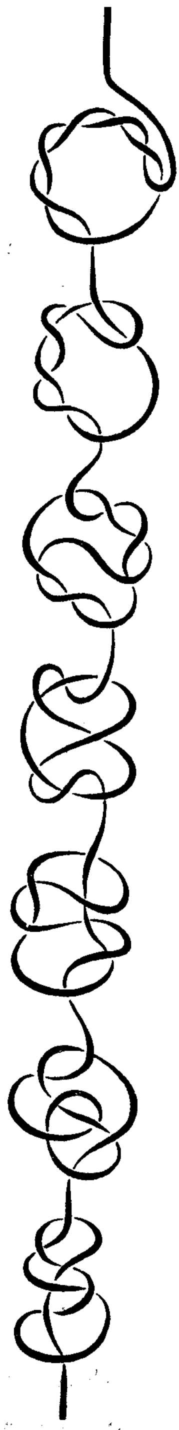
Richard H. Crowell
Ralph H. Fox

Introduction to Knot Theory

Springer—Verlag
New York Heidelberg Berlin
World Publishing Corporation Beijing China

Richard H. Crowell
Ralph H. Fox

Introduction to Knot Theory



Springer-Verlag
New York Heidelberg Berlin
World Publishing Corporation Beijing,China

Preface to the Springer Edition

This book was written as an introductory text for a one-semester course and, as such, it is far from a comprehensive reference work. Its lack of completeness is now more apparent than ever since, like most branches of mathematics, knot theory has expanded enormously during the last fifteen years. The book could certainly be rewritten by including more material and also by introducing topics in a more elegant and up-to-date style. Accomplishing these objectives would be extremely worthwhile. However, a significant revision of the original work along these lines, as opposed to writing a new book, would probably be a mistake. As inspired by its senior author, the late Ralph H. Fox, this book achieves qualities of effectiveness, brevity, elementary character, and unity. These characteristics would be jeopardized, if not lost, in a major revision. As a result, the book is being republished unchanged, except for minor corrections. The most important of these occurs in Chapter III, where the old sections 2 and 3 have been interchanged and somewhat modified. The original proof of the theorem that a group is free if and only if it is isomorphic to $F[\mathscr{A}]$ for some alphabet $\mathscr{A}$ contained an error, which has been corrected using the fact that equivalent reduced words are equal.

I would like to include a tribute to Ralph Fox, who has been called the father of modern knot theory. He was indisputably a first-rate mathematician of international stature. More importantly, he was a great human being. His students and other friends respected him, and they also loved him. This edition of the book is dedicated to his memory.

Richard H. Crowell

Dartmouth College
1977

Preface

Knot theory is a kind of geometry, and one whose appeal is very direct because the objects studied are perceivable and tangible in everyday physical space. It is a meeting ground of such diverse branches of mathematics as group theory, matrix theory, number theory, algebraic geometry, and differential geometry, to name some of the more prominent ones. It had its origins in the mathematical theory of electricity and in primitive atomic physics, and there are hints today of new applications in certain branches of chemistry.[1] The outlines of the modern topological theory were worked out by Dehn, Alexander, Reidemeister, and Seifert almost thirty years ago. As a subfield of topology, knot theory forms the core of a wide range of problems dealing with the position of one manifold imbedded within another.

This book, which is an elaboration of a series of lectures given by Fox at Haverford College while a Philips Visitor there in the spring of 1956, is an attempt to make the subject accessible to everyone. Primarily it is a textbook for a course at the junior-senior level, but we believe that it can be used with profit also by graduate students. Because the algebra required is not the familiar commutative algebra, a disproportionate amount of the book is given over to necessary algebraic preliminaries. However, this is all to the good because the study of noncommutativity is not only essential for the development of knot theory but is itself an important and not overcultivated field. Perhaps the most fascinating aspect of knot theory is the interplay between geometry and this noncommutative algebra.

For the past thirty years Kurt Reidemeister's Ergebnisse publication *Knotentheorie* has been virtually the only book on the subject. During that time many important advances have been made, and moreover the combinatorial point of view that dominates *Knotentheorie* has generally given way to a strictly topological approach. Accordingly, we have emphasized the topological invariance of the theory throughout.

There is no doubt whatever in our minds but that the subject centers around the concepts: *knot group*, *Alexander matrix*, *covering space*, and our presentation is faithful to this point of view. We regret that, in the interest of keeping the material at as elementary a level as possible, we did not introduce and make systematic use of covering space theory. However, had we done so, this book would have become much longer, more difficult, and

[1] H.L. Frisch and E. Wasserman, "Chemical Topology," *J. Am. Chem. Soc.*, **83** (1961) 3789–3795

presumably also more expensive. *For the mathematician with some maturity,* for example one who has finished studying this book, a survey of this central core of the subject may be found in Fox's "A quick trip through knot theory" (1962).[1]

The bibliography, although not complete, is comprehensive far beyond the needs of an introductory text. This is partly because the field is in dire need of such a bibliography and partly because we expect that our book will be of use to even sophisticated mathematicians well beyond their student days. To make this bibliography as useful as possible, we have included a *guide to the literature*.

Finally, we thank the many mathematicians who had a hand in reading and criticizing the manuscript at the various stages of its development. In particular, we mention Lee Neuwirth, J. van Buskirk, and R. J. Aumann, and two Dartmouth undergraduates, Seth Zimmerman and Peter Rosmarin. We are also grateful to David S. Cochran for his assistance in updating the bibliography for the third printing of this book.

---

[1] See Bibliography

Contents

Chapter VI · Presentation of a Knot Group

Chapter VII · The Free Calculus and the Elementary Ideals

Chapter VIII · The Knot Polynomials

Chapter IX · Characteristic Properties of the Knot Polynomials

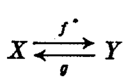
Prerequisites

For an intelligent reading of this book a knowledge of the elements of modern algebra and point-set topology is sufficient. Specifically, we shall assume that the reader is familiar with the concept of a function (or mapping) and the attendant notions of domain, range, image, inverse image, one-one, onto, composition, restriction, and inclusion mapping; with the concepts of equivalence relation and equivalence class; with the definition and elementary properties of open set, closed set, neighborhood, closure, interior, induced topology, Cartesian product, continuous mapping, homeomorphism, compactness, connectedness, open cover(ing), and the Euclidean n-dimensional space R^n; and with the definition and basic properties of homomorphism, automorphism, kernel, image, groups, normal subgroups, quotient groups, rings, (two-sided) ideals, permutation groups, determinants, and matrices. These matters are dealt with in many standard textbooks. We may, for example, refer the reader to A. H. Wallace, *An Introduction to Algebraic Topology* (Pergamon Press, 1957), Chapters I, II, and III, and to G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Revised Edition (The Macmillan Co., New York, 1953), Chapters III, §§1-3, 7, 8; VI, §§4-8, 11-14; VII, §5; X, §§1, 2; XIII, §§1-4. Some of these concepts are also defined in the index.

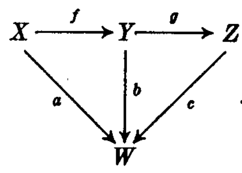
In Appendix I an additional requirement is a knowledge of differential and integral calculus.

The usual set theoretic symbols $\in$, $\subset$, $\supset$, $=$, $\cup$, $\cap$, and $-$ are used. For the inclusion symbol we follow the common convention: $A \subset B$ means that $p \in B$ whenever $p \in A$. For the image and inverse image of A under f we write either fA and $f^{-1}A$, or $f(A)$ and $f^{-1}(A)$. For the restriction of f to A we write $f \mid A$, and for the composition of two mappings $f\colon X \to Y$ and $g\colon Y \to Z$ we write gf.

When several mappings connecting several sets are to be considered at the same time, it is convenient to display them in a (mapping) diagram, such as



or



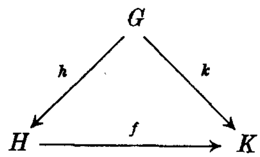
.

If each element in each set displayed in a diagram has at most one image element in any given set of the diagram, the diagram is said to be ***consistent.***

Thus the first diagram is consistent if and only if $gf = 1$ and $fg = 1$, and the second diagram is consistent if and only if $bf = a$ and $cg = b$ (and hence $cgf = a$).

The reader should note the following "diagram-filling" lemma, the proof of which is straightforward.

If $h\colon G \to H$ and $k\colon G \to K$ are homomorphisms and h is onto, there exists a (necessarily unique) homomorphism $f\colon H \to K$ making the diagram



consistent if and only if the kernel of h is contained in the kernel of k.

CHAPTER I

Knots and Knot Types

1. Definition of a knot. Almost everyone is familiar with at least the simplest of the common knots, e.g., the overhand knot, Figure 1, and the figure-eight knot, Figure 2. A little experimenting with a piece of rope will convince anyone that these two knots are different: one cannot be transformed into the other without passing a loop over one of the ends, i.e.,without "tying" or "untying." Nevertheless, failure to change the figure-eight into the overhand by hours of patient twisting is no proof that it can't be done. The problem that we shall consider is the problem of showing mathematically that these knots (and many others) are distinct from one another.

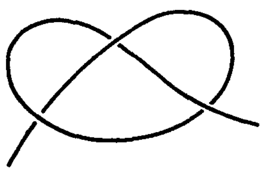


Figure 1

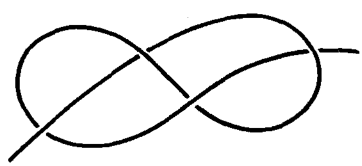


Figure 2

Mathematics never proves anything about anything except mathematics, and a piece of rope is a physical object and not a mathematical one. So before worrying about proofs, we must have a mathematical definition of what a knot is and another mathematical definition of when two knots are to be considered the same. This problem of formulating a mathematical model arises whenever one applies mathematics to a physical situation. The definitions should define mathematical objects that approximate the physical objects under consideration as closely as possible. The model may be good or bad according as the correspondence between mathematics and reality is good or bad. There is, however, no way to prove (in the mathematical sense, and it is probably only in this sense that the word has a precise meaning) that the mathematical definitions describe the physical situation exactly.

Obviously, the figure-eight knot can be transformed into the overhand knot by tying and untying—in fact all knots are equivalent if this operation is allowed. Thus tying and untying must be prohibited either in the definition of when two knots are to be considered the same or from the beginning in the very definition of what a knot is. The latter course is easier and is the one we shall adopt. Essentially, we must get rid of the ends. One way would be to prolong the ends to infinity; but a simpler method is to splice them together. Accordingly, we shall consider a knot to be a subset of 3-dimensional space which is homeomorphic to a circle. The formal definition is: K is a *knot* if there exists a homeomorphism of the unit circle C into 3-dimensional space R^3 whose image is K. By the circle C is meant the set of points (x,y) in the plane R^2 which satisfy the equation $x^2 + y^2 = 1$.

The overhand knot and the figure-eight knot are now pictured as in Figure 3 and Figure 4. Actually, in this form the overhand knot is often called the *clover-leaf knot*. Another common name for this knot is the *trefoil*. The figure-eight knot has been called both the *four-knot* and *Listing's knot*.

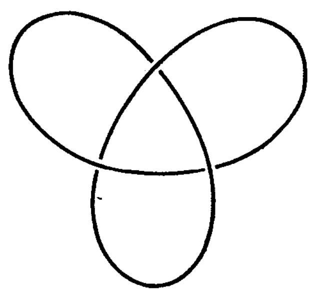


Figure 3

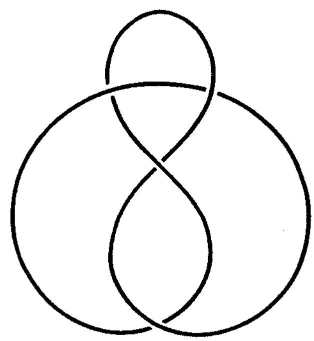


Figure 4

We next consider the question of when two knots K_1 and K_2 are to be considered the same. Notice, first of all, that this is not a question of whether or not K_1 and K_2 are homeomorphic. They are both homeomorphic to the unit circle and, consequently, to each other. The property of being knotted is not an intrinsic topological property of the space consisting of the points of the knot, but is rather a characteristic of the way in which that space is imbedded in R^3. Knot theory is a part of 3-dimensional topology and not of 1-dimensional topology. If a piece of rope in one position is twisted into another, the deformation does indeed determine a one-one correspondence between the points of the two positions, and since cutting the rope is not allowed, the correspondence is bicontinuous. In addition, it is natural to think of the motion of the rope as accompanied by a motion of the surrounding air molecules which thus determines a bicontinuous permutation of the points of space. This picture suggests the definition: Knots K_1 and K_2 are *equivalent* if there exists a homeomorphism of R^3 onto itself which maps K_1 onto K_2.

It is a triviality that the relation of knot equivalence is a true equivalence relation. Equivalent knots are said to be of the same *type*, and each equivalence class of knots is a *knot type*. Those knots equivalent to the unknotted circle $x^2 + y^2 = 1$, $z = 0$, are called *trivial* and constitute the *trivial type*.[1] Similarly, the type of the clover-leaf knot, or of the figure-eight knot is defined as the equivalence class of some particular representative knot. The informal statement that the clover-leaf knot and the figure-eight knot are different is rigorously expressed by saying that they belong to distinct knot types.

2. Tame versus wild knots. A *polygonal knot* is one which is the union of a finite number of closed straight-line segments called *edges*, whose endpoints are the *vertices* of the knot. A knot is *tame* if it is equivalent to a polygonal knot; otherwise it is *wild*. This distinction is of fundamental importance. In fact, most of the knot theory developed in this book is applicable (as it stands) only to tame knots. The principal invariants of knot type, namely, the elementary ideals and the knot polynomials, are not necessarily defined for a wild knot. Moreover, their evaluation is based on finding a polygonal representative to start with. The discovery that knot theory is largely confined to the study of polygonal knots may come as a surprise—especially to the reader who approaches the subject fresh from the abstract generality of point-set topology. It is natural to ask what kinds of knots other than polygonal are tame. A partial answer is given by the following theorem.

(2.1) *If a knot parametrized by arc length is of class* C^1 (*i.e., is continuously differentiable*), *then it is tame.*

A proof is given in Appendix I. It is complicated but straightforward, and it uses nothing beyond the standard techniques of advanced calculus. More explicitly, the assumptions on K are that it is rectifiable and given as the image of a vector-valued function $p(s) = (x(s), y(s), z(s))$ of arc length s with continuous first derivatives. Thus, every sufficiently smooth knot is tame.

It is by no means obvious that there exist any wild knots. For example, no knot that lies in a plane is wild. Although the study of wild knots is a corner of knot theory outside the scope of this book, Figure 5 gives an example of a knot known to be wild.[2] This knot is a remarkable curve. Except for the fact that the number of loops increases without limit while their size decreases without limit (as is indicated in the figure by the dotted square about p), the

---

[1] Any knot which lies in a plane is necessarily trivial. This is a well-known and deep theorem of plane topology. See A. H. Newman, *Elements of the Topology of Plane Sets of Points*, Second edition (Cambridge University Press, Cambridge, 1951), p. 173.

[2] R. H. Fox, "A Remarkable Simple Closed Curve," *Annals of Mathematics*, Vol. 50 (1949), pp. 264, 265.

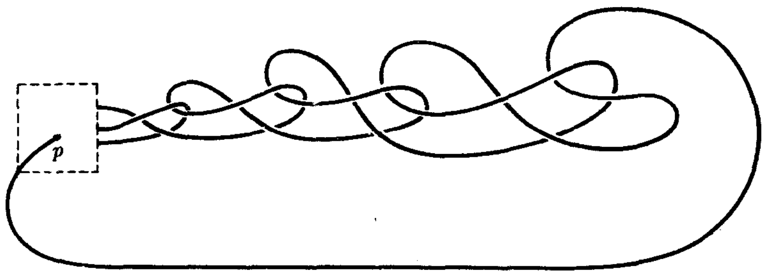


Figure 5

knot could obviously be untied. Notice also that, except at the single point p, it is as smooth and differentiable as we like.

3. Knot projections. A knot K is usually specified by a projection; for example, Figure 3 and Figure 4 show projected images of the clover-leaf knot and the figure-eight knot, respectively. Consider the parallel projection

$$\mathscr{P}\colon R^3 \to R^3$$

defined by $\mathscr{P}(x,y,z) = (x,y,0)$. A point p of the image $\mathscr{P}K$ is called a *multiple point* if the inverse image $\mathscr{P}^{-1}p$ contains more than one point of K. The *order* of $p \in \mathscr{P}K$ is the cardinality of $(\mathscr{P}^{-1}p) \cap K$. Thus, a *double point* is a multiple point of order 2, a *triple point* is one of order 3, and so on. Multiple points of infinite order can also occur. In general, the image $\mathscr{P}K$ may be quite complicated in the number and kinds of multiple points present. It is possible, however, that K is equivalent to another knot whose projected image is fairly simple. For a polygonal knot, the criterion for being fairly simple is that the knot be in what is called regular position. The definition is as follows: a polygonal knot K is in *regular position* if: (i) the only multiple points of K are double points, and there are only a finite number of them; (ii) no double point is the image of any vertex of K. The second condition insures that every double point depicts a genuine crossing, as in Figure 6a. The sort of double point shown in Figure 6b is prohibited.

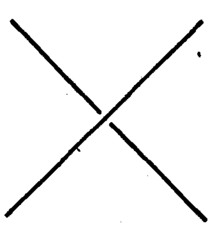


Figure 6a

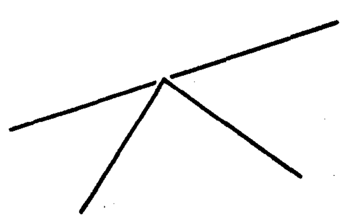


Figure 6b

Each double point of the projected image of a polygonal knot in regular position is the image of two points of the knot. The one with the larger z-coordinate is called an *overcrossing*, and the other is the corresponding *undercrossing*.

(3.1) *Any polygonal knot K is equivalent under an arbitrarily small rotation of R^3 to a polygonal knot in regular position.*

Proof. The geometric idea is to hold K fixed and move the projection. Every bundle (or pencil) of parallel lines in R^3 determines a unique parallel projection of R^3 onto the plane through the origin perpendicular to the bundle. We shall assume the obvious extension of the above definition of regular position so that it makes sense to ask whether or not K is in regular position with respect to any parallel projection. It is convenient to consider R^3 as a subset[3] of a real projective 3-space P^3. Then, to every parallel projection we associate the point of intersection of any line parallel to the direction of projection with the projective plane P^2 at infinity. This correspondence is clearly one-one and onto. Let Q be the set of all points of P^2 corresponding to projections with respect to which K is not in regular position. We shall show that Q is nowhere dense in P^2. It then follows that there is a projection $\mathscr{P}_0$ with respect to which K is in regular position and which is arbitrarily close to the original projection $\mathscr{P}$ along the z-axis. Any rotation of R^3 which transforms the line $\mathscr{P}_0^{-1}(0,0,0)$ into the z-axis will suffice to complete the proof.

In order to prove that Q is nowhere dense in P^2, consider first the set of all straight lines which join a vertex of K to an edge of K. These intersect P^2 in a finite number of straight-line segments whose union we denote by Q_1. Any projection corresponding to a point of $P^2 - Q_1$ must obviously satisfy condition (ii) of the definition of regular position. Furthermore, it can have at most a finite number of multiple points, no one of which is of infinite order. It remains to show that multiple points of order $n \geq 3$ can be avoided, and this is done as follows. Consider any three mutually skew straight lines, each of which contains an edge of K. The locus of all straight lines which intersect these three is a quadric surface which intersects P^2 in a conic section. (See the reference in the preceding footnote.) Set Q_2 equal to the union of all such conics. Obviously, there are only a finite number of them. Furthermore, the image of K under any projection which corresponds to some point of $P^2 - (Q_1 \cup Q_2)$ has no multiple points of order $n \geq 3$. We have shown that

$$P^2 - (Q_1 \cup Q_2) \subset P^2 - Q.$$

Thus Q is a subset of $Q_1 \cup Q_2$, which is nowhere dense in P^2. This completes the proof of (3.1).

[3] For an account of the concepts used in this proof, see O. Veblen and J. W. Young, *Projective Geometry* (Ginn and Company, Boston, Massachusetts, 1910), Vol. 1 pp. 11, 299, 301.

Thus, every tame knot is equivalent to a polygonal knot in regular position. This fact is the starting point for calculating the basic invariants by which different knot types are distinguished.

4. Isotopy type, amphicheiral and invertible knots. This section is not a prerequisite for the subsequent development of knot theory in this book. The contents are nonetheless important and worth reading even on the first time through.

Our definition of knot type was motivated by the example of a rope in motion from one position in space to another and accompanied by a displacement of the surrounding air molecules. The resulting definition of equivalence of knots abstracted from this example represents a simplification of the physical situation, in that no account is taken of the motion during the transition from the initial to the final position. A more elaborate construction, which does model the motion, is described in the definition of the isotopy type of a knot. An *isotopic deformation* of a topological space X is a family of homeomorphisms h_t, $0 \leq t \leq 1$, of X onto itself such that h_0 is the identity, i.e., $h_0(p) = p$ for all p in X, and the function H defined by $H(t,p) = h_t(p)$ is simultaneously continuous in t and p. This is a special case of the general definition of a deformation which will be studied in Chapter V. Knots K_1 and K_2 are said to belong to the same *isotopy type* if there exists an isotopic deformation $\{h_t\}$ of R^3 such that $h_1K_1 = K_2$. The letter t is intentionally chosen to suggest time. Thus, for a fixed point $p \in R^3$, the point $h_t(p)$ traces out, so to speak, the path of the molecule originally at p during the motion of the rope from its initial position at K_1 to K_2.

Obviously, if knots K_1 and K_2 belong to the same isotopy type, they are equivalent. The converse, however, is false. The following discussion of orientation serves to illustrate the difference between the two definitions.

Every homeomorphism h of R^3 onto itself is either *orientation preserving* or *orientation reversing*. Although a rigorous treatment of this concept is usually given by homology theory,[4] the intuitive idea is simple. The homeomorphism h preserves orientation if the image of every right (left)-hand screw is again a right (left)-hand screw; it reverses orientation if the image of every right (left)-hand screw is a left (right)-hand screw. The reason that there is no other possibility is that, owing to the continuity of h, the set of points of R^3 at which the twist of a screw is preserved by h is an open set and the same is true of the set of points at which the twist is reversed. Since h is a homeo-

[4] A homeomorphism k of the n-sphere S^n, $n \geq 1$, onto itself is *orientation preserving* or *reversing* according as the isomorphism $k_*: H_n(S^n) \to H_n(S^n)$ is or is not the identity. Let $S^n = R^n \cup \{\infty\}$ be the one point compactification of the real Cartesian n-space R^n. Any homeomorphism h of R^n onto itself has a unique extention to a homeomorphism k of $S^n = R^n \cup \{\infty\}$ onto itself defined by $k \mid R^n = h$ and $k(\infty) = \infty$. Then, h is *orientation preserving* or *reversing* according as k is orientation preserving or reversing.

morphism, every point of R^3 belongs to one of these two disjoint sets; and since R^3 is connected, it follows that one of the two sets is empty. The composition of homeomorphisms follows the usual rule of parity:

h_1	h_2	h_1h_2
preserving	preserving	preserving
reversing	preserving	reversing
preserving	reversing	reversing
reversing	reversing	preserving

Obviously, the identity mapping is orientation preserving. On the other hand, the reflection $(x,y,z) \to (x,y,-z)$ is orientation reversing. If h is a linear transformation, it is orientation preserving or reversing according as its determinant is positive or negative. Similarly, if both h and its inverse are C^1 differentiable at every point of R^3, then h preserves or reverses orientation according as its Jacobian is everywhere positive or everywhere negative.

Consider an isotopic deformation $\{h_t\}$ of R^3. The fact that the identity is orientation preserving combined with the continuity of $H(t,p) = h_t(p)$, suggests that h_t is orientation preserving for every t in the interval $0 \leq t \leq 1$. This is true.[5] As a result, we have that a necessary condition for two knots to be of the same isotopy type is that there exist an orientation preserving homeomorphism of R^3 on itself which maps one knot onto the other.

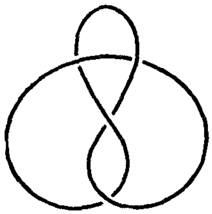
A knot K is said to be *amphicheiral* if there exists an orientation reversing homeomorphism h of R^3 onto itself such that $hK = K$. An equivalent formulation of the definition, which is more appealing geometrically, is provided by the following lemma. By the *mirror image* of a knot K we shall mean the image of K under the reflection $\mathscr{R}$ defined by $(x,y,z) \to (x,y,-z)$. Then,

(4.1) *A knot K is amphicheiral if and only if there exists an orientation preserving homeomorphism of R^3 onto itself which maps K onto its mirror image.*

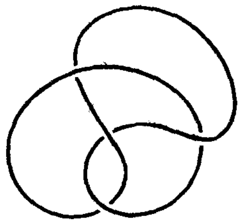
Proof. If K is amphicheiral, the composition $\mathscr{R}h$ is orientation preserving and maps K onto its mirror image. Conversely, if h' is an orientation preserving homeomorphism of R^3 onto itself which maps K onto its mirror image, the composition $\mathscr{R}h'$ is orientation reversing and $(\mathscr{R}h')K = K$.

It is not hard to show that the figure-eight knot is amphicheiral. The experimental approach is the best; a rope which has been tied as a figure-eight and then spliced is quite easily twisted into its mirror image. The operation is illustrated in Figure 7. On the other hand, the clover-leaf knot is not amphi-

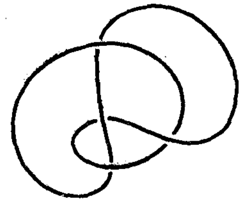
[5] Any isotopic deformation $\{h_t\}$, $0 \leq t \leq 1$, of the Cartesian n-space R^n definitely possesses a unique extension to an isotopic deformation $\{k_t\}$, $0 \leq t \leq 1$, of the n-sphere S^n, i.e., $k_t \mid R^n = h_t$, and $k_t(\infty) = \infty$. For each t, the homeomorphism k_t is homotopic to the identity, and so the induced isomorphism $(k_t)_*$ on $H_n(S^n)$ is the identity. It follows that h_t is orientation preserving for all t in $0 \leq t \leq 1$. (See also footnote 4.)



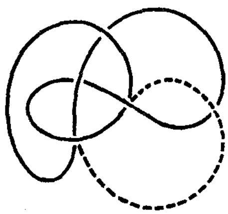
(1)



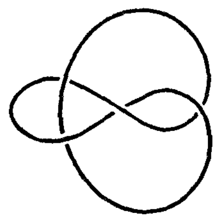
(2)



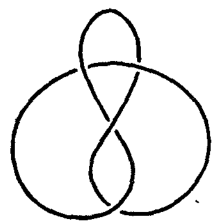
(3)



(4)



(5)



(6)

cheiral. In this case, experimenting with a piece of rope accomplishes nothing except possibly to convince the skeptic that the question is nontrivial. Actually, to prove that the clover-leaf is not amphicheiral is hard and requires fairly advanced techniques of knot theory. Assuming this result, however, we have that the clover-leaf knot and its mirror image are equivalent but not of the same isotopy type.

It is natural to ask whether or not every orientation preserving homeomorphism f of R^3 onto itself is realizable by an isotopic deformation, i.e., given f, does there exist $\{h_t\}$, $0 \leq t \leq 1$, such that $f = h_1$? If the answer were no, we would have a third kind of knot type. This question is not an easy one. The answer is, however, yes.[6]

Just as every homeomorphism of R^3 onto itself either preserves or reverses orientation, so does every homeomorphism f of a knot K onto itself. The geometric interpretation is analogous to, and simpler than, the situation in 3-dimensional space. Having prescribed a direction on the knot, f preserves or reverses orientation according as the order of points of K is preserved or reversed under f. A knot K is called *invertible* if there exists an orientation preserving homeomorphism h of R^3 onto itself such that the restriction $h \mid K$ is an orientation reversing homeomorphism of K onto itself. Both the clover-

[6] G. M. Fisher, "On the Group of all Homeomorphisms of a Manifold," *Transactions of the American Mathematical Society*, Vol. 97 (1960), pp. 193–212.