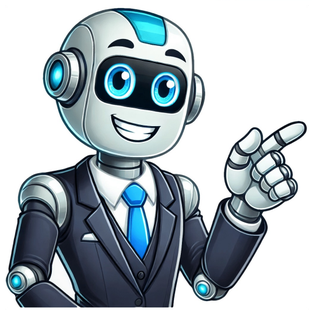
Click to prove you're human



Mark Belan/Quanta Magazine Illuminating basic science and math research through public service journalism. More about us Quanta Magazine is committed to in-depth, accurate journalism that serves the public interest. Each article braids the complexities of science with the malleable art of storytelling and is meticulously reported, edited and fact-checked. Launched and funded by the Simons Foundation, Quanta is editorially independent — our articles do not reflect or represent the views of the foundation. More about us The Quanta Newsletter Get highlights of the most important news delivered to your email inbox Knot theory is a branch of mathematics that studies mathematical knots, which are closed loops in three-dimensional space. These knots can be formed by intertwining and looping segments of a rope without cutting or overlapping. Knot theory has applications in various fields, including physics, chemistry, biology, and computer science. For instance, it's used in understanding the behavior of polymers, the structure of DNA molecules, and the study of fluid dynamics. What is Knot Theory?Knot theory is a branch of mathematics which deals with the study of knots. Mathematical knots are well-defined as closed curves embedded in three-dimensional space. These knots can be defined as knots that contain no loops of the string with intersection points, and they are constructed by looping a single string. In mathematics, knots can be classified according to properties of some kind, for instance, the number of crossing points and their symmetry. The knots are then used to study and categorize these complex structures in an attempt to predict their behavior and other characteristics. Mathematical Definition of KnotIn mathematics, a knot is formally defined as a closed curve embedded in three-dimensional space. More precisely, a knot is a smooth, continuous mapping of the circle S1 into three-dimensional Euclidean space R3, such that the image of the circle is a compact, connected, and oriented one-dimensional manifold without boundary. This means that the curve does not intersect itself, except possibly at its endpoints, and it does not have loose ends. Mathematically, a knot is often represented as a smooth function: K : S1 → R3 Where each point on the circle S1 corresponds to a unique point in three-dimensional space R3, tracing out the entire knot curve. Examples of KnotsSome common examples of knots used in everyday life include: Trivial Knot (Unknot): This knot is simply a closed loop with no twists or crossings. It's the simplest knot, often represented as 01 or 0. Trefoil Knot: One of the simplest non-trivial knots, it has the form of three loops twisted around each other. It's represented as 31 in knot notation. Figure-Eight Knot: This knot resembles the number eight, formed by two loops crossing each other. It's represented as 41 in knot notation. Square Knot: Also known as the granny knot or reef knot, it's formed by joining two ends of a rope after each end has been passed around the other. This knot is widely used in daily life, for instance, in tying shoelaces. Types of KnotsKnots in knot theory are classified based on the complexity of the knot and the pattern of the knot as stated below. Some categories of knots are simple knots; the complex ones, prime knots and composite knots. Simple KnotsA simple knot is an elementary type of knot with no complex twist patterns and no loops within the knot. They do not require to be cut to untie and can be regarded as the basis for forming more intricate knots. Some examples of simple knots include: Trefoil Knot: This knot is the simplest non-trivial knot.Figure-Eight Knot: Also known as the "reef knot" or "savoy knot," this knot resembles the number eight (∞) and is formed by crossing the rope over itself twice.Overhand Knot: This is one of the simplest knots, formed by passing one end of a rope over itself and through the loop created.Square Knot (Reef Knot): The square knot is used to join two ropes of similar size. It consists of two overhand knots, one tied in the opposite direction of the other, resulting in a flat, symmetrical knot.Complex KnotsComplex knots are knots that are not simple, meaning they can be untangled or deformed into a simple loop without cutting. They exhibit a higher level of complexity compared to simple knots and require detailed analysis to understand their topology. Some example of complex knots are: Figure-Eight Follow-Through Knot(41 - 41): This knot is more complex than the simple figure-eight knot. It is formed by passing a rope through a figure-eight knot, resulting in a knot that is more secure and often used in climbing and sailing.Sheet Bend (52): Also known as the weaver's knot, the sheet bend is used to join two ropes of different diameters. It is formed by passing one rope through a loop formed by the other rope and then wrapping it around both ropes.Double Fisherman's Knot(51 - 51: This knot is used to join two ropes together. It is formed by tying two overhand knots in each rope and then cinching them tight against each other.Monkey's Fist (63): This knot is tied at the end of a rope and is often used as a weight or as a decorative knot. It is formed by wrapping the rope around itself multiple times and then weaving it back through the wraps.Prime KnotsAn example of an important type of knot is the prime knot which is defined as a knot that cannot be simplified and reduced any further. They are necessary in the study of knots and are the building blocks from which all other knots can be formed. Some examples of Prime Knots are: Cinquefoil Knot (51): This knot has five crossings and is also a prime knot.Stevedore Knot (61): With six crossings, the stevedore knot is another example of a prime knot.Overhand Knot (31 - 31): Despite its simplicity, the overhand knot is considered prime because it cannot be further decomposed into simpler knots.Composite KnotsComposite knots are those knots that are made by the use of more than one simple knot. These knots show elaborate patterns of interconnection and reveal ways in which separate knot elements relate to each other. Studying composite knots helps in understanding the relationships between different knot configurations. Knot DiagramsA knot diagram is simply a picture of a knot on paper, it is a flat picture of a knot. It is most beneficial within the field of knot theory, as it enables the viewer to observe knot structures in detail. A knot diagram has a set of crossings over or under a sequence of arcs, which represents segments of the rope between crossings. In a knot diagram, the crossings are depicted by a gap in a particular or its representation meaning that somewhere the rope is actually going around the knot. For instance, the position of the break will inform the status of whether the rope crosses over "crossing over break" or crosses under "crossing under break". Knots are easily distinguishable employing this notation to describe and depict the structure of the given knot. Knot diagrams are essential in knot theory because they enable the application of various mathematical techniques to study knots. For example, the Reidemeister moves are a set of three basic transformations that can be applied to a knot diagram to simplify it without changing the underlying knot. These moves include: Reidemeister Move 1: A twist in the rope that does not change the number of crossings.Reidemeister Move 2: A move that eliminates two consecutive crossings of the same type.Reidemeister Move 3: A move that slides a strand over another strand, changing the order of crossings.By applying these moves, mathematicians can simplify a knot diagram, making it easier to analyze and classify the knot. Applications of Knot TheoryThe applications of Knot theory is as follows: DNA Analysis: Knot theory is used in understanding the topology of DNA strands and analyzing genetic structures.Robotics: Knot theory is applied in robot motion planning and control to avoid entanglements and optimize movements.Chemistry: Knot theory helps in studying molecular structures and analyzing the entanglement of polymers.Physics: Knot theory is used in quantum field theory and particle physics to study topological defects and solitons.Computer Science: Knot theory finds applications in cryptography and error correction algorithms for data transmission.ConclusionKnot theory is a specialized branch of mathematics that focuses on studying various 'knots' – complex formations. As a result of the study of knots' properties, classifications, and uses, mathematicians obtain significant knowledge about topology and geometry. Knot theory is not exclusive to mere mathematics, but it has found incorporation in various other disciplines and technologies. Read More, Number TheorySet TheoryGame Theory Study of mathematical knots Examples of different knots including the trivial knot (top left) and the trefoil knot (below it) A knot diagram of the trefoil knot, the simplest non-trivial knot In topology, knot theory is the study of mathematical knots. While inspired by knots which appear in daily life, such as those in shoelaces and rope, a mathematical knot differs in that the ends are joined so it cannot be undone, the simplest knot being a ring (or "unknot"). In mathematical language, a knot is an embedding of a circle in 3-dimensional Euclidean space, E 3 {\displaystyle \mathbb {E} ^{3}} . Two mathematical knots are equivalent if one can be transformed into the other via a deformation of R 3 {\displaystyle \mathbb {R} ^{3}} upon itself (known as an ambient isotopy); these transformations correspond to manipulations of a knotted string that do not involve cutting it or passing it through itself. Knots can be described in various ways. Using different description methods, there may be more than one description of the same knot. For example, a common method of describing a knot is a planar diagram called a knot diagram, in which any knot can be drawn in many different ways. Therefore, a fundamental problem in knot theory is determining when two descriptions represent the same knot. A complete algorithmic solution to this problem exists, which has unknown complexity.[1] In practice, knots are often distinguished using a knot invariant, a "quantity" which is the same when computed from different descriptions of a knot. Important invariants include knot polynomials, knot groups, and hyperbolic invariants. The original motivation for the founders of knot theory was to create a table of knots and links, which are knots of several components entangled with each other. More than six billion knots and links have been tabulated since the beginnings of knot theory in the 19th century. To gain further insight, mathematicians have generalized the knot concept in several ways. Knots can be considered in other three-dimensional spaces and objects other than circles can be used; see knot (mathematics). For example, a higher-dimensional knot is an n-dimensional sphere embedded in (n+2)-dimensional Euclidean space. Main article: History of knot theory Intricate Celtic knotwork in the 1200-year-old Book of Kells Archaeologists have discovered that knot tying dates back to prehistoric times. Besides their uses such as recording information and tying objects together, knots have interested humans for their aesthetics and spiritual symbolism. Knots appear in various forms of Chinese artwork dating from several centuries BC (see Chinese knotting). The endless knot appears in Tibetan Buddhism, while the Borromean rings have made repeated appearances in different cultures, often representing strength in unity. The Celtic monks who created the Book of Kells lavished entire pages with intricate Celtic knotwork. The first knot tabulator, Peter Guthrie Tait A mathematical theory of knots was first developed in 1771 by Alexandre-Théophile Vandermonde who explicitly noted the importance of topological features when discussing the properties of knots related to the geometry of position. Mathematical studies of knots began in the 19th century with Carl Friedrich Gauss, who defined the linking integral (Silver 2006). In the 1860s, Lord Kelvin's theory that atoms were knots in the aether led to Peter Guthrie Tait's creation of the first knot tables for complete classification. Tait, in 1885, published a table of knots with up to ten crossings, and what came to be known as the Tait conjectures. This record motivated the early knot theorists, but knot theory eventually became part of the emerging subject of topology. These topologists in the early part of the 20th century—Max Dehn, J. W. Alexander, and others—studied knots from the point of view of the knot group and invariants from homology theory such as the Alexander polynomial. This would be the main approach to knot theory until a series of breakthroughs transformed the subject. In the late 1970s, William Thurston introduced hyperbolic geometry into the study of knots with the hyperbolization theorem. Many knots were shown to be hyperbolic knots, enabling the use of geometry in defining new, powerful knot invariants. The discovery of the Jones polynomial by Vaughan Jones in 1984 (Sossinsky 2002, pp. 71–89), and subsequent contributions from Edward Witten, Maxim Kontsevich, and others, revealed deep connections between knot theory and mathematical methods in statistical mechanics and quantum field theory. A plethora of knot invariants have been invented since then, utilizing sophisticated tools such as quantum groups and Floer homology. In the last several decades of the 20th century, scientists became interested in studying physical knots in order to understand knotting phenomena in DNA and other polymers. Knot theory can be used to determine if a molecule is chiral (has a "handedness") or not (Simon 1986). Tangles, strings with both ends fixed in place, have been effectively used in studying the action of topoisomerase on DNA (Flapan 2000). Knot theory may be crucial in the construction of quantum computers, through the model of topological quantum computation (Collins 2006). On the left, the unknot, and a knot equivalent to it. It can be more difficult to determine whether complex knots, such as the one on the right, are equivalent to the unknot. A knot is created by beginning with a one-dimensional line segment, wrapping it around itself arbitrarily, and then fusing its two free ends together to form a closed loop (Adams 2004) (Sossinsky 2002). Simply, we can say a knot K {\displaystyle K} is a "simple closed curve" (see Curve) — that is: a "nearly" injective and continuous function K : [0 , 1] → R 3 {\displaystyle K\colon [0,1]\to \mathbb {R} ^{3}} , with the only "non-injectivity" being K (0) = K (1) {\displaystyle K(0)=K(1)} . Topologists consider knots and other entanglements such as links and braids to be equivalent if the knot can be pushed about smoothly, without intersecting itself, to coincide with another knot. The idea of knot equivalence is to give a precise definition of when two knots should be considered the same even when positioned quite differently in space. A formal mathematical definition is that two knots K 1 , K 2 {\displaystyle K_{1},K_{2}} are equivalent if there is an orientation-preserving homeomorphism h : R 3 → R 3 {\displaystyle h\colon \mathbb {R} ^{3}\to \mathbb {R} ^{3}} with h (K 1) = K 2 {\displaystyle h(K_{1})=K_{2}} . What this definition of knot equivalence means is that two knots are equivalent when there is a continuous family of homeomorphisms { h t : R 3 → R 3 f o r 0 ≤ t ≤ 1 } {\displaystyle \{h_{t}:\mathbb {R} ^{3}\rightarrow \mathbb {R} ^{3}\ \mathrm {for} \ 0\leq t\leq 1\}} of space onto itself, such that the last one of them carries the first knot onto the second knot. (In detail: Two knots K 1 {\displaystyle K_{1}} and K 2 {\displaystyle K_{2}} are equivalent if there exists a continuous mapping H : R 3 × [0 , 1] → R 3 {\displaystyle H:\mathbb {R} ^{3}\times [0,1]\rightarrow \mathbb {R} ^{3}} such that a) for each t ∈ [0 , 1] {\displaystyle t\in [0,1]} the mapping taking x ∈ R 3 {\displaystyle x\in \mathbb {R} ^{3}} to H (x , t) ∈ R 3 {\displaystyle H(x,t)\in \mathbb {R} ^{3}} is a homeomorphism of R 3 {\displaystyle \mathbb {R} ^{3}} onto itself; b) H (x , 0) = x {\displaystyle H(x,0)=x} for all x ∈ R 3 {\displaystyle x\in \mathbb {R} ^{3}} ; and c) H (K 1 , 1) = K 2 {\displaystyle H(K_{1},1)=K_{2}} . Such a function H {\displaystyle H} is known as an ambient isotopy.) These two notions of knot equivalence agree exactly about which knots are equivalent: Two knots that are equivalent under the orientation-preserving homeomorphism definition are also equivalent under the ambient isotopy definition, because any orientation-preserving homeomorphisms of R 3 {\displaystyle \mathbb {R} ^{3}} to itself is the final stage of an ambient isotopy starting from the identity. Conversely, two knots equivalent under the ambient isotopy definition are also equivalent under the orientation-preserving homeomorphism definition, because the t = 1 {\displaystyle t=1} (final) stage of the ambient isotopy must be an orientation-preserving homeomorphism carrying one knot to the other. The basic problem of knot theory, the recognition problem, is determining the equivalence of two knots. Algorithms exist to solve this problem, with the first given by Wolfgang Haken in the late 1960s (Hass 1998). Nonetheless, these algorithms can be extremely time-consuming, and a major issue in the theory is to understand how hard this problem really is (Hass 1998). The special case of recognizing the unknot, called the unknotting problem, is of particular interest (Hoste 2005). In February 2021 Marc Lackenby announced a new unknot recognition algorithm that runs in quasi-polynomial time.[2] Tenfold Knottiness, plate IX, from Peter Guthrie Tait's article "On Knots", 1884 A useful way to visualise and manipulate knots is to project the knot onto a plane—think of the knot casting a shadow on the wall. A small change in the direction of projection will ensure that it is one-to-one except at the double points, called crossings, where the "shadow" of the knot crosses itself once transversely (Rolfsen 1976). At each crossing, to be able to recreate the original knot, the over-strand must be distinguished from the under-strand. This is often done by creating a break in the strand going underneath. The resulting diagram is an immersed plane curve with the additional data of which strand is over and which is under at each crossing. (These diagrams are called knot diagrams when they represent a knot and link diagrams when they represent a link.) Analogously, knotted surfaces in 4-space can be related to immersed surfaces in 3-space. A reduced diagram is a knot diagram in which there are no reducible crossings (also nugatory or removable crossings), or in which all of the reducible crossings have been removed.[3][4] A petal projection is a type of projection in which, instead of forming double points, all strands of the knot meet at a single crossing point, connected to it by loops forming non-nested "petals".[5] Main article: Reidemeister move In 1927, working with this diagrammatic form of knots, J. W. Alexander and Garland Baird Briggs, and independently Kurt Reidemeister, demonstrated that two knot diagrams belonging to the same knot can be related by a sequence of three kinds of moves on the diagram, shown below. These operations, now called the Reidemeister moves, are: Twist and untwist in either direction.Move one strand completely over another.Move a strand completely over or under a crossing. Reidemeister moves Type I Type II Type III The proof that diagrams of equivalent knots are connected by Reidemeister moves relies on an analysis of what happens under the planar projection of the movement taking one knot to another. The movement can be arranged so that almost all of the time the projection will be a knot diagram, except at finitely many times when an "event" or "catastrophe" occurs, such as when more than two strands cross at a point or multiple strands become tangent at a point. A close inspection will show that complicated events can be eliminated, leaving only the simplest events: (1) a "kink" forming or being straightened out; (2) two strands becoming tangent at a point and passing through; and (3) three strands crossing at a point. These are precisely the Reidemeister moves (Sossinsky 2002, ch. 3) (Lickorish 1997, ch. 1). A 3D print depicting the complement of the figure eight knotby François Guéritaud, Saul Schleimer, and Henry Segerman Main article: Knot invariant A knot invariant is a "quantity" that is the same for equivalent knots (Adams 2004) (Lickorish 1997) (Rolfsen 1976). For example, if the invariant is computed from a knot diagram, it should give the same value for two knot diagrams representing equivalent knots. An invariant may take the same value on two different knots, so by itself may be incapable of distinguishing all knots. An elementary invariant is tricolorability. "Classical" knot invariants include the knot group, which is the fundamental group of the knot complement, and the Alexander polynomial, which can be computed from the Alexander invariant, a module constructed from the infinite cyclic cover of the knot complement (Lickorish 1997)(Rolfsen 1976). In the late 20th century, invariants such as "quantum" knot polynomials, Vassiliev invariants and hyperbolic invariants were discovered. These aforementioned invariants are only the tip of the iceberg of modern knot theory. Main article: Knot polynomial A knot polynomial is a knot invariant that is a polynomial. Well-known examples include the Jones polynomial, the Alexander polynomial, and the Kauffman polynomial. A variant of the Alexander polynomial, the Alexander–Conway polynomial, is a polynomial in the variable z with integer coefficients (Lickorish 1997). The Alexander–Conway polynomial is actually defined in terms of links, which consist of one or more knots entangled with each other. The concepts explained above for knots, e.g. diagrams and Reidemeister moves, also hold for links. Consider an oriented link diagram, i.e. one in which every component of the link has a preferred direction indicated by an arrow. For a given crossing of the diagram, let L + , L − , L 0 {\displaystyle L_{+},L_{-},L_{0}} be the oriented link diagrams resulting from changing the diagram as indicated in the figure: The original diagram might be either L + {\displaystyle L_{+}} or L − {\displaystyle L_{-}} , depending on the chosen crossing's configuration. Then the Alexander–Conway polynomial, C (z) {\displaystyle C(z)} , is recursively defined according to the rules: C (O) = 1 {\displaystyle C(O)=1} (where O {\displaystyle O} is any diagram of the unknot) C (L +) = C (L −) + z C (L 0) . {\displaystyle C(L_{+})=C(L_{-})+zC(L_{0}).} The second rule is what is often referred to as a skein relation. To check that these rules give an invariant of an oriented link, one should determine that the polynomial does not change under the three Reidemeister moves. Many important knot polynomials can be defined in this way. The following is an example of a typical computation using a skein relation. It computes the Alexander–Conway polynomial of the trefoil knot. The yellow patches indicate where the relation is applied. C() = C() + z C() gives the unknot and the Hopf link. Applying the relation to the Hopf link where indicated, C() = C() + z C() gives a link deformable to one with 0 crossings (it is actually the unlink of two components) and an unknot. The unlink takes a bit of sneakiness: C() = C() + z C() which implies that C(unlink of two components) = 0, since the first two polynomials are of the unknot and thus equal. Putting all this together will show: C (t r e f o i l) = 1 + z (0 + z) = 1 + z 2 {\displaystyle C(\mathrm {trefoil})=1+z(0+z)=1+z^{2}} Since the Alexander–Conway polynomial is a knot invariant, this shows that the trefoil is not equivalent to the unknot. So the trefoil really is "knotted". The left-handed trefoil knot. The right-handed trefoil knot. Actually, there are two trefoil knots, called the right and left-handed trefoils, which are mirror images of each other (take a diagram of the trefoil given above and change each crossing to the other way to get the mirror image). These are not equivalent to each other, meaning that they are not amphichiral. This was shown by Max Dehn, before the invention of knot polynomials, using group theoretical methods (Dehn 1914). But the Alexander–Conway polynomial of each kind of trefoil will be the same, as can be seen by going through the computation above with the mirror image. The Jones polynomial can in fact distinguish between the left- and right-handed trefoil knots (Lickorish 1997). William Thurston proved many knots are hyperbolic knots, meaning that the knot complement (i.e., the set of points of 3-space not on the knot) admits a geometric structure, in particular that of hyperbolic geometry. The hyperbolic structure depends only on the knot so any quantity computed from the hyperbolic structure is then a knot invariant (Adams 2004). The Borromean rings are a link with the property that removing one ring unlinks the others.SnapPea's cusp view: the Borromean rings complement from the perspective of an inhabitant living near the red component. Geometry lets us visualize what the inside of a knot or link complement looks like by imagining light rays as traveling along the geodesics of the geometry. An example is provided by the picture of the complement of the Borromean rings. The inhabitant of this link complement is viewing the space from near the red component. The balls in the picture are views of horoball neighborhoods of the link. By thickening the link in a standard way, the horoball neighborhoods of the link components are obtained. Even though the boundary of a neighborhood is a torus, when viewed from inside the link complement, it looks like a sphere. Each link component shows up as infinitely many spheres (of one color) as there are infinitely many light rays from the observer to the link component. The fundamental parallelogram (which is indicated in the picture), tiles both vertically and horizontally and shows how to extend the pattern of spheres infinitely. This pattern, the horoball pattern, is itself a useful invariant. Other hyperbolic invariants include the shape of the fundamental parallelogram, length of shortest geodesic, and volume. Modern knot and link tabulation efforts have utilized these invariants effectively. Fast computers and clever methods of obtaining these invariants make calculating these invariants, in practice, a simple task (Adams, Hildebrand & Weeks 1991). A knot in three dimensions can be untied when placed in four-dimensional space. This is done by changing crossings. Suppose one strand is behind another as seen from a chosen point. Lift it into the fourth dimension, so there is no obstacle (the front strand having no component there); then slide it forward, and drop it back, now in front. Analogies for the plane would be lifting a string up off the surface, or removing a dot from inside a circle. In fact, in four dimensions, any non-intersecting closed loop of one-dimensional string is equivalent to an unknot. First "push" the loop into a three-dimensional subspace, which is always possible, though technical to explain. Four-dimensional space occurs in classical knot theory, however, and an important topic is the study of slice knots and ribbon knots. A notorious open problem asks whether every slice knot is also ribbon. Since a knot can be considered topologically a 1-dimensional sphere, the next generalization is to consider a two-dimensional sphere (S 2 {\displaystyle \mathbb {S} ^{2}}) embedded in 4-dimensional Euclidean space (R 4 {\displaystyle \mathbb {R} ^{4}}). Such an embedding is knotted if there is no homeomorphism of R 4 {\displaystyle \mathbb {R} ^{4}} onto itself taking the embedded 2-sphere to the standard "round" embedding of the 2-sphere. Suspended knots and spun knots are two typical families of such 2-sphere knots. The mathematical technique called "general position" implies that for a given n-sphere in m-dimensional Euclidean space, if m is large enough (depending on n), the sphere should be unknotted. In general, piecewise-linear n-spheres form knots only in (n + 2)-dimensional space (Zeeman 1963), although this is no longer a requirement for smoothly knotted spheres. In fact, there are smoothly knotted (4 k − 1) {\displaystyle (4k-1)} -spheres in 6k-dimensional space; e.g., there is a smoothly knotted 3-sphere in R 6 {\displaystyle \mathbb {R} ^{6}} (Haefliger 1962) (Levine 1965). Thus the codimension of a smooth knot can be arbitrarily large when not fixing the dimension of the knotted sphere; however, any smooth k-sphere embedded in R n {\displaystyle \mathbb {R} ^{n}} with 2 n − 3 k − 3 > 0 {\displaystyle 2n-3k-3>0} is unknotted. The notion of a knot has further generalisations in mathematics, see: Knot (mathematics), isotopy classification of embeddings. Every knot in the n-sphere S n {\displaystyle \mathbb {S} ^{n}} is the link of a real-algebraic set with isolated singularity in R n + 1 {\displaystyle \mathbb {R} ^{n+1}} (Akbulut & King 1981). An n-knot is a single S n {\displaystyle \mathbb {S} ^{n}} embedded in R m {\displaystyle \mathbb {R} ^{m}} . An n-link consists of k-copies of S n {\displaystyle \mathbb {S} ^{n}} embedded in R m {\displaystyle \mathbb {R} ^{m}} , where k is a natural number. Both the m = n + 2 {\displaystyle m=n+2} and the m > n + 2 {\displaystyle m>n+2} cases are well studied, and so is the n > 1 {\displaystyle n>1} case.[6][7] Main article: Knot sum Adding two knots Two knots can be added by cutting both knots and joining the pairs of ends. The operation is called the knot sum, or sometimes the connected sum or composition of two knots. This can be formally defined as follows (Adams 2004): consider a planar projection of each knot and suppose these projections are disjoint. Find a rectangle in the plane where one pair of opposite sides are arcs along each knot while the rest of the rectangle is disjoint from the knots. Form a new knot by deleting the first pair of opposite sides and adjoining the other pair of opposite sides. The resulting knot is a sum of the original knots. Depending on how this is done, two different knots (but no more) may result. This ambiguity in the sum can be eliminated regarding the knots as oriented, i.e. having a preferred direction of travel along the knot, and requiring the arcs of the knots in the sum are oriented consistently with the oriented boundary of the rectangle. The knot sum of oriented knots is commutative and associative. A knot is prime if it is non-trivial and cannot be written as the knot sum of two non-trivial knots. A knot that can be written as such a sum is composite. There is a prime decomposition for knots, analogous to prime and composite numbers (Schubert 1949). For oriented knots, this decomposition is also unique. Higher-dimensional knots can also be added but there are some differences. While you cannot form the unknot in three dimensions by adding two non-trivial knots, you can in higher dimensions, at least when one considers smooth knots in codimension at least 3. Knots can also be constructed using the circuit topology approach. This is done by combining basic units called soft contacts using five operations (Parallel, Series, Cross, Concerted, and Sub).[8][9] The approach is applicable to open chains as well and can also be extended to include the so-called hard contacts. See also: List of prime knots and Knot tabulation A table of prime knots up to seven crossings. The knots are labeled with Alexander–Briggs notation Traditionally, knots have been catalogued in terms of crossing number. Knot tables generally include only prime knots, and only one entry for a knot and its mirror image (even if they are different) (Hoste, Thistlethwaite & Weeks 1998). The number of nontrivial knots of a given crossing number increases rapidly, making tabulation computationally difficult (Hoste 2005, p. 20). Tabulation efforts have succeeded in enumerating over 6 billion knots and links (Hoste 2005, p. 28). The sequence of the number of prime knots of a given crossing number, up to crossing number 16, is 0, 0, 1, 1, 2, 3, 7, 21, 49, 165, 552, 2176, 9988, 46972, 253293, 1388705... (sequence A002863 in the OEIS). While exponential upper and lower bounds for this sequence are known, it has not been proven that this sequence is strictly increasing (Adams 2004). The first knot tables by Tait, Little, and Kirkman used knot diagrams, although Tait also used a precursor to the Dowker notation. Different notations have been invented for knots which allow more efficient tabulation (Hoste 2005). The early tables attempted to list all knots of at most 10 crossings, and all alternating knots of 11 crossings (Hoste, Thistlethwaite & Weeks 1998). The development of knot theory due to Alexander, Reidemeister, Seifert, and others eased the task of verification and tables of knots up to and including 9 crossings were published by Alexander–Briggs and Reidemeister in the late 1920s. The first major verification of this work was done in the 1960s by John Horton Conway, who not only developed a new notation but also the Alexander–Conway polynomial (Conway 1970) (Doll & Hoste 1991). This verified the list of knots of at most 11 crossings and a new list of links up to 10 crossings. Conway found a number of omissions but only one duplication in the Tait–Little tables; however he missed the duplicates called the Perko pair, which would only be noticed in 1974 by Kenneth Perko (Perko 1974). This famous error would propagate when Dale Rolfsen added a knot table in his influential text, based on Conway's work. Conway's 1970 paper on knot theory also contains a typographical duplication on its non-alternating 11-crossing knots page and omits 4 examples — 2 previously listed in D. Lombardero's 1968 Princeton senior thesis and 2 more subsequently discovered by Alain Caudron. [see Perko (1982), Primality of certain knots, Topology Proceedings] Less famous is the duplicate in his 10 crossing link table: 2.-2.-20.20 is the mirror of 8*-20:-20. [See Perko (2016), Historical highlights of non-cyclic knot theory, J. Knot Theory Ramifications]. In the late 1990s Hoste, Thistlethwaite, and Weeks tabulated all the knots through 16 crossings (Hoste, Thistlethwaite & Weeks 1998). In 2003 Rankin, Flint, and Schermann, tabulated the alternating knots through 22 crossings (Hoste 2005). In 2020 Burton tabulated all prime knots with up to 19 crossings (Burton 2020). This is the most traditional notation, due to the 1927 paper of James W. Alexander and Garland B. Briggs and later extended by Dale Rolfsen in his knot table (see image above and List of prime knots). The notation simply organizes knots by their crossing number. One writes the crossing number with a subscript to denote its order amongst all knots with that crossing number. This order is arbitrary and so has no special significance (though in each number of crossings the twist knot comes after the torus knot). Links are written by the crossing number with a superscript to denote the number of components and a subscript to denote its order within the links with the same number of components and crossings. Thus the trefoil knot is notated 31 and the Hopf link is 221. Alexander–Briggs names in the range 10162 to 10166 are ambiguous, due to the discovery of the Perko pair in Charles Newton Little's original and subsequent knot tables, and differences in approach to correcting this error in knot tables and other publications created after this point.[10] Main article: Dowker–Thistlethwaite notation A knot diagram with crossings labelled for a Dowker sequence The Dowker–Thistlethwaite notation, also called the Dowker notation or code, for a knot is a finite sequence of even integers. The numbers are generated by following the knot and marking the crossings with consecutive integers. Since each crossing is visited twice, this creates a pairing of even integers with odd integers. An appropriate sign is given to indicate over and undercrossing. For example, in this figure the knot diagram has crossings labelled with the pairs (1,6) (3,−12) (5,2) (7,8) (9,−4) and (11,−10). The Dowker–Thistlethwaite notation for this labelling is the sequence: 6, −12, 2, 8, −4, −10. A knot diagram has more than one possible Dowker notation, and there is a well-understood ambiguity when reconstructing a knot from a Dowker–Thistlethwaite notation. Main article: Conway notation (knot theory) The Conway notation for knots and links, named after John Horton Conway, is based on the theory of tangles (Conway 1970). The advantage of this notation is that it reflects some properties of the knot or link. The notation describes how to construct a particular link diagram of the link. Start with a basic polyhedron, a 4-valent connected planar graph with no digon regions. Such a polyhedron is denoted first by the number of vertices then a number of asterisks which determine the polyhedron's position on a list of basic polyhedra. For example, 10** denotes the second 10-vertex polyhedron on Conway's list. Each vertex then has an algebraic tangle substituted into it (each vertex is oriented so there is no arbitrary choice in substitution). Each such tangle has a notation consisting of numbers and + or − signs. An example is 1*2 −3 2. The 1* denotes the only 1-vertex basic polyhedron. The 2 −3 2 is a sequence describing the continued fraction associated to a rational tangle. One inserts this tangle at the vertex of the basic polyhedron 1*. A more complicated example is 8*3.1.2 0.1.1.1.1.1 Here again 8* refers to a basic polyhedron with 8 vertices. The periods separate the notation for each tangle. Any link admits such a description, and it is clear this is a very compact notation even for very large crossing number. There are some further shorthands usually used. The last example is usually written 8*3:2 0, where the ones are omitted and kept the number of dots excepting the dots at the end. For an algebraic knot such as in the first example, 1* is often omitted. Conway's pioneering paper on the subject lists up to 10-vertex basic polyhedra of which he uses to tabulate links, which have become standard for those links. For a further listing of higher vertex polyhedra, there are nonstandard choices available. Main article: Gauss code Gauss code, similar to the Dowker–Thistlethwaite notation, represents a knot with a sequence of integers. However, rather than every crossing being represented by two different numbers, crossings are labeled with only one number. When the crossing is an overcrossing, a positive number is listed. At an undercrossing, a negative number. For example, the trefoil knot in Gauss code can be given as: 1,−2,3,−1,2,−3 Gauss code is limited in its ability to identify knots. This problem is partially addressed with by the extended Gauss code. Arithmetic rope Circuit topology Lamp cord trick Legendrian submanifolds and knots List of knot theory topics Molecular knot Necktie § Knots Quantum topology Ribbon theory Adams, Colin (2004), The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots, American Mathematical Society, ISBN 978-0-8218-3678-1 Adams, Colin; Crawford, Thomas; DeMeo, Benjamin; Landry, Michael; Lin, Alex Tong; Montee, MurphyKate; Park, Seojung; Venkatesh, Saraswathi; Yhee, Farrah (2015), "Knot projections with a single multi-crossing", Journal of Knot Theory and Its Ramifications, 24 (3): 1550011, 30, arXiv:1208.5742, doi:10.1142/S021821651550011X, MR 3342136, S2CID 119320887 Adams, Colin; Hildebrand, Martin; Weeks, Jeffrey (1991), "Hyperbolic invariants of knots and links", Transactions of the American Mathematical Society, 326 (1): 1–56, doi:10.1090/s0002-9947-1991-0994161-2, JSTOR 2001854 Akbulut, Selman; King, Henry C. (1981), "All knots are algebraic", Comment. Math. Helv., 56 (3): 339–351, doi:10.1007/BF02566217, S2CID 120218312 Bar-Natan, Dror (1995), "On the Vassiliev knot invariants", Topology, 34 (2): 423–472, doi:10.1016/0040-9383(95)93237-2 Burton, Benjamin A. (2020). "The Next 350 Million Knots". 36th International Symposium on Computational Geometry (SoCG 2020). Leibniz Int. Proc. Inform. Vol. 164. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. pp. 25:1–25:17. doi:10.4230/LIPIcs.SoCG.2020.25. Collins, Graham (April 2006), "Computing with Quantum Knots", Scientific American, 294 (4): 56–63, Bibcode:2006SciAm.294d..56C, doi:10.1038/scientificamerican0406-56, PMID 16596880 Dehn, Max (1914), "Die beiden Kleeblattschlingen", Mathematische Annalen, 75 (3): 402–413, doi:10.1007/BF01563732, S2CID 120452571 Conway, John H. (1970), "An enumeration of knots and links, and some of their algebraic properties", Computational Problems in Abstract Algebra, Pergamon, pp. 329–358, doi:10.1016/B978-0-08-012975-4.50034-5, ISBN 978-0-08-012975-4 Doll, Helmut; Hoste, Jim (1991), "A tabulation of oriented links. With microfiche supplement", Math. Comp., 57 (196): 747–761, Bibcode:1991MaCom..57..747D, doi:10.1090/S0025-5718-1991-1094946-4 Flapan, Erica (2000), When topology meets chemistry: A topological look at molecular chirality, Outlook, Cambridge University Press, ISBN 978-0-521-66254-3 Haefliger, André (1962), "Knotted (4k − 1)-spheres in 6k-space", Annals of Mathematics, Second Series, 75 (3): 452–466, doi:10.2307/1970208, JSTOR 1970208 Haken, Wolfgang (1962), "Über das Homöomorphieproblem der 3-Mannigfaltigkeiten. I", Mathematische Zeitschrift, 80: 89–120, doi:10.1007/BF01162369, ISSN 0025-5874, MR 0160196 Hass, Joel (1998), "Algorithms for recognizing knots and 3-manifolds", Chaos, Solitons and Fractals, 9 (4–5): 569–581, arXiv:math/9712269, Bibcode:1998CSF.....9..569H, doi:10.1016/S0960-0779(97)00109-4, S2CID 7381505 Hoste, Jim; Thistlethwaite, Morwen; Weeks, Jeffrey (1998), "The First 1,701,935 Knots", Math. Intelligencer, 20 (4): 33–48, doi:10.1007/BF03025227, S2CID 18027155 Hoste, Jim (2005). "The Enumeration and Classification of Knots and Links". Handbook of Knot Theory. pp. 209–232. doi:10.1016/B978-044451452-3/50006-X. ISBN 978-0-444-51452-3. Levine, Jerome (1965), "A classification of differentiable knots", Annals of Mathematics, Second Series, 1982 (1): 15–50, doi:10.2307/1970561, JSTOR 1970561 Kontsevich, M. (1993). "Vassiliev's knot invariants". I. M. Gelfand Seminar. ADVSOV. Vol. 16. pp. 137–150. doi:10.1090/advsov/016.2/04. ISBN 978-0-8218-4117-4. Lickorish, W. B. Raymond (1997), An Introduction to Knot Theory, Graduate Texts in Mathematics, vol. 175, Springer-Verlag, doi:10.1007/978-1-4612-0691-0, ISBN 978-0-387-98254-0, S2CID 122824389 Perko, Kenneth (1974), "On the classification of knots", Proceedings of the American Mathematical Society, 45 (2): 262–6, doi:10.2307/2040074, JSTOR 2040074 Rolfsen, Dale (1976), Knots and Links, Mathematics Lecture Series, vol. 7, Berkeley, California: Publish or Perish, ISBN 978-0-914098-16-4, MR 0515288 Schubert, Horst (1949). Die eindeutige Zerlegbarkeit eines Knotens in Primknoten. doi:10.1007/978-3-642-45813-2. ISBN 978-3-540-01419-5. {{cite book}}: ISBN / Date incompatibility (help) Silver, Daniel (2006). "Knot Theory's Odd Origins". American Scientist. 94 (2): 158. doi:10.1511/2006.2.158. Simon, Jonathan (1986), "Topological chirality of certain molecules", Topology, 25 (2): 229–235, doi:10.1016/0040-9383(86)90041-8 Sossinsky, Alexei (2002), Knots, mathematics with a twist, Harvard University Press, ISBN 978-0-674-00944-8 Turaev, Vladimir G. (2016). Quantum Invariants of Knots and 3-Manifolds. doi:10.1515/9783110

the lines radiating out from the center of the sphere. The intersection of a plane that contains the normal with the surface will form a curve that is called a normal section, and the curvature of this curve is the normal curvature. For most points on most surfaces, different sections will have different curvatures; the maximum and minimum values of these are called the principal curvatures. Any closed surface will have at least four points called umbilical points. At an umbilic all the sectional curvatures are equal; in particular the principal curvatures are equal. Umbilical points can be thought of as the points where the surface is closely approximated by a sphere. For the sphere the curvatures of all normal sections are equal, so every point is an umbilic. The sphere and plane are the only surfaces with this property. The sphere does not have a surface of centers. For a given normal section exists a circle of curvature that equals the sectional curvature, is tangent to the surface, and the center lines of which lie along on the normal line. For example, the two centers corresponding to the maximum and minimum sectional curvatures are called the focal points, and the set of all such centers forms the focal surface. For most surfaces the focal surface forms two sheets that are each a surface and meet at umbilical points. Several cases are special: * For channel surfaces one sheet forms a curve and the other sheet is a surface * For cones, cylinders, tori and cyclides both sheets form curves. * For the sphere the center of every osculating circle is at the center of the sphere and the focal surface forms a single point. This property is unique to the sphere. All geodesics of the sphere are closed curves. Geodesics are curves on a surface that give the shortest distance between two points. They are a generalization of the concept of a straight line in the plane. For the sphere the geodesics are great circles. Many other surfaces share this property. Of all the solids having a given volume, the sphere is the one with the smallest surface area; of all solids having a given surface area, the sphere is the one having the greatest volume. It follows from isoperimetric inequality. These properties define the sphere uniquely and can be seen in soap bubbles: a soap bubble will enclose a fixed volume, and surface tension minimizes its surface area for that volume. A freely floating soap bubble therefore approximates a sphere (though such external forces as gravity will slightly distort the bubble's shape). It can also be seen in planets and stars where gravity minimizes surface area for large celestial bodies. The sphere has the smallest total mean curvature among all convex solids with a given surface area. The mean curvature is the average of the two principal curvatures, which is constant because the two principal curvatures are constant at all points of the sphere. The sphere has constant mean curvature. The sphere is the only embedded surface that lacks boundary or singularities with constant positive mean curvature. Other such immersed surfaces as minimal surfaces have constant mean curvature. The sphere has constant positive Gaussian curvature. Gaussian curvature is the product of the two principal curvatures. It is an intrinsic property that can be determined by measuring length and angles and is independent of how the surface is embedded in space. Hence, bending a surface will not alter the Gaussian curvature, and other surfaces with constant positive Gaussian curvature can be obtained by cutting a small slit in the sphere and bending it. All these other surfaces would have boundaries, and the sphere is the only surface that lacks a boundary with constant, positive Gaussian curvature. The pseudosphere is an example of a surface with constant negative Gaussian curvature. The sphere is transformed into itself by a three-parameter family of rigid motions. Rotating around any axis a unit sphere at the origin will map the sphere onto itself. Any rotation about a line through the origin can be expressed as a combination of rotations around the three-coordinate axis (see Euler angles). Therefore, a three-parameter family of rotations exists such that each rotation transforms the sphere onto itself; this family is the rotation group SO(3). The plane is the only other surface with a three-parameter family of transformations (translations along the x- and y-axes and rotations around the origin). Circular cylinders are the only surfaces with two-parameter families of rigid motions and the surfaces of revolution and helicoids are the only surfaces with a one-parameter family. Main article: Spherical geometry Great circle on a sphere The basic elements of Euclidean plane geometry are points and lines. On the sphere, points are defined in the usual sense. The analogue of the "line" is the geodesic, which is a great circle; the defining characteristic of a great circle is that the plane containing all its points also passes through the center of the sphere. Measuring by arc length shows that the shortest path between two points lying on the sphere is the shorter segment of the great circle that includes the points. Many theorems from classical geometry hold true for spherical geometry as well, but not all do because the sphere fails to satisfy some of classical geometry's postulates, including the parallel postulate. In spherical trigonometry, angles are defined between great circles. Spherical trigonometry differs from ordinary trigonometry in many respects. For example, the sum of the interior angles of a spherical triangle always exceeds 180 degrees. Also, any two similar spherical triangles are congruent. Any pair of points on a sphere that lie on a straight line through the sphere's center (i.e., the diameter) are called antipodal points – on the sphere, the distance between them is exactly half the length of the circumference.[note 2] Any other (i.e., not antipodal) pair of distinct points on a sphere lie on a unique great circle, segment it into one minor (i.e., shorter) and one major (i.e., longer) arc, and have the minor arc's length be the shortest distance between them on the sphere.[note 3] Spherical geometry is a form of elliptic geometry, which together with hyperbolic geometry makes up non-Euclidean geometry. The sphere is a smooth surface with constant Gaussian curvature at each point equal to 1/r2.[9] As per Gauss's Theorema Egregium, this curvature is independent of the sphere's embedding in 3-dimensional space. Also following from Gauss, a sphere cannot be mapped to a plane while maintaining both areas and angles. Therefore, any map projection introduces some form of distortion. A sphere of radius r has area element d A = r 2 sin θ d θ d φ {\displaystyle dA=r^{2}\sin \theta \,d\theta \,d\varphi } . This can be found from the volume element in spherical coordinates with r held constant.[9] A sphere of any radius centered at zero is an integral surface of the following differential form: x d x + y d y + z d z = 0. {\displaystyle x\,dx+y\,dy+z\,dz=0.} This equation reflects that the position vector and tangent plane at a point are always orthogonal to each other. Furthermore, the outward-facing normal vector is equal to the position vector scaled by 1/r. In Riemannian geometry, the filling area conjecture states that the hemisphere is the optimal (least area) isometric filling of the Riemannian circle. Remarkably, it is possible to turn an ordinary sphere inside out in a three-dimensional space with possible self-intersections but without creating any creases, in a process called sphere eversion. The antipodal quotient of the sphere is the surface called the real projective plane, which can also be thought of as the Northern Hemisphere with antipodal points of the equator identified. Plane section of a sphere: one circle Coaxial intersection of a sphere and a cylinder: two circles Main article: Circle of a sphere Circles on the sphere are, like circles in the plane, made up of all points a certain distance from a fixed point on the sphere. The intersection of a sphere and a plane is a circle, a point, or empty.[18] Great circles are the intersection of the sphere with a plane passing through the center of a sphere: others are called small circles. More complicated surfaces may intersect a sphere in circles, too: the intersection of a sphere with a surface of revolution whose axis contains the center of the sphere (are coaxial) consists of circles and/or points if not empty. For example, the diagram to the right shows the intersection of a sphere and a cylinder, which consists of two circles. If the cylinder radius were that of the sphere, the intersection would be a single circle. If the cylinder radius were larger than that of the sphere, the intersection would be empty. Main article: Rhumb line Loxodrome In navigation, a loxodrome or rhumb line is a path whose bearing, the angle between its tangent and due North, is constant. Loxodromes project to straight lines under the Mercator projection. Two special cases are the meridians which are aligned directly North–South and parallels which are aligned directly East–West. For any other bearing, a loxodrome spirals infinitely around each pole. For the Earth modeled as a sphere, or for a general sphere given a spherical coordinate system, such a loxodrome is a kind of spherical spiral.[19] Main article: Clélie Clelia spiral with c = 8 Another kind of spherical spiral is the Clelia curve, for which the longitude (or azimuth) φ {\displaystyle \varphi } and the colatitude (or polar angle) θ {\displaystyle \theta } are in a linear relationship, φ = c θ {\displaystyle \varphi =c\theta } . Clelia curves project to straight lines under the equirectangular projection. Viviani's curve (c = 1 {\displaystyle c=1}) is a special case. Clelia curves approximate the ground track of satellites in polar orbit. Main article: Spherical conic The analog of a conic section on the sphere is a spherical conic, a quartic curve which can be defined in several equivalent ways. The intersection of a sphere with a quadratic cone whose vertex is the sphere center The intersection of a sphere with an elliptic or hyperbolic cylinder whose axis passes through the sphere center The locus of points whose sum or difference of great-circle distances from a pair of foci is a constant Many theorems relating to planar conic sections also extend to spherical conics. General intersection sphere-cylinder If a sphere is intersected by another surface, there may be more complicated spherical curves. Example sphere–cylinder Main article: Sphere-cylinder intersection The intersection of the sphere with equation x 2 + y 2 + z 2 = r 2 {\displaystyle \;x^{2}+y^{2}+z^{2}=r^{2}\;} and the cylinder with equation (y − y 0) 2 + z 2 = a 2 , y 0 ≠ 0 {\displaystyle \;(y-y_{0})^{2}+z^{2}=a^{2},\;y_{0}eq 0\;} is not just one or two circles. It is the solution of the non-linear system of equations x 2 + y 2 + z 2 − r 2 = 0 {\displaystyle x^{2}+y^{2}+z^{2}-r^{2}=0} (y − y 0) 2 + z 2 − a 2 = 0 . {\displaystyle (y-y_{0})^{2}+z^{2}-a^{2}=0\ .} (see implicit curve and the diagram) An ellipsoid is a sphere that has been stretched or compressed in one or more directions. More exactly, it is the image of a sphere under an affine transformation. An ellipsoid bears the same relationship to the sphere that an ellipse does to a circle. Main article: n-sphere Spheres can be generalized to spaces of any number of dimensions. For any natural number n, an n-sphere, often denoted Sn, is the set of points in (n + 1)-dimensional Euclidean space that are at a fixed distance r from a central point of that space, where r is, as before, a positive real number. In particular: S0: a 0-sphere consists of two discrete points, −r and r S1: a 1-sphere is a circle of radius r S2: a 2-sphere is an ordinary sphere S3: a 3-sphere is a sphere in 4-dimensional Euclidean space. Spheres for n > 2 are sometimes called hyperspheres. The n-sphere of unit radius centered at the origin is denoted Sn and is often referred to as "the" n-sphere. The ordinary sphere is a 2-sphere, because it is a 2-dimensional surface which is embedded in 3-dimensional space. In topology, the n-sphere is an example of a compact topological manifold without boundary. A topological sphere need not be smooth; if it is smooth, it need not be diffeomorphic to the Euclidean sphere (an exotic sphere). The sphere is the inverse image of a one-point set under the continuous function ‖x‖, so it is closed; Sn is also bounded, so it is compact by the Heine–Borel theorem. Main article: Metric space More generally, in a metric space (E,d), the sphere of center x and radius r > 0 is the set of points y such that d(x,y) = r. If the center is a distinguished point that is considered to be the origin of E, as in a normed space, it is not mentioned in the definition and notation. The same applies for the radius if it is taken to equal one, as in the case of a unit sphere. Unlike a ball, even a large sphere may be an empty set. For example, in Zn with Euclidean metric, a sphere of radius r is nonempty only if r2 can be written as sum of n squares of integers. An octahedron is a sphere in taxicab geometry, and a cube is a sphere in geometry using the Chebyshev distance. The geometry of the sphere was studied by the Greeks. Euclid's Elements defines the sphere in book XI, discusses various properties of the sphere in book XII, and shows how to inscribe the five regular polyhedra within a sphere in book XIII. Euclid does not include the area and volume of a sphere, only a theorem that the volume of a sphere varies as the third power of its diameter, probably due to Eudoxus of Cnidus. The volume and area formulas were first determined in Archimedes's On the Sphere and Cylinder by the method of exhaustion. Zenodorus was the first to state that, for a given surface area, the sphere is the solid of maximum volume.[3] Archimedes wrote about the problem of dividing a sphere into segments whose volumes are in a given ratio, but did not solve it. A solution by means of the parabola and hyperbola was given by Dionysodorus.[20] A similar problem – to construct a segment equal in volume to a given segment, and in surface to another segment – was solved later by al-Quhi.[3] An image of one of the most accurate human-made spheres, as it refracts the image of Einstein in the background. This sphere was a fused quartz gyroscope for the Gravity Probe B experiment, and differs in shape from a perfect sphere by no more than 40 atoms (less than 10 nm) of thickness. It was announced on 1 July 2008 that Australian scientists had created even more nearly perfect spheres, accurate to 0.3 nm, as part of an international hunt to find a new global standard kilogram.[21] Deck of playing cards illustrating engineering instruments, England, 1702. King of spades: Spheres See also: Ball (mathematics) § Regions Hemisphere Spherical cap Spherical lune Spherical polygon Spherical sector Spherical segment Spherical wedge Spherical zone 3-sphere Affine sphere Alexander horned sphere Celestial spheres Curvature Directional statistics Dyson sphere Gauss map Hand with Reflecting Sphere, M.C. Escher self-portrait drawing illustrating reflection and the optical properties of a mirror sphere Hoberman sphere Homology sphere Homotopy groups of spheres Homotopy sphere Lenart Sphere Napkin ring problem Orb (optics) Pseudosphere Riemann sphere Solid angle Sphere packing Spherical coordinates Spherical cow Spherical helix, tangent indicatrix of a curve of constant precession Spherical polyhedron Sphericity Tennis ball theorem Volume-equivalent radius Zoll sphere ^ r is being considered as a variable in this computation. ^ It does not matter which direction is chosen, the distance is the sphere's radius × π. ^ The distance between two non-distinct points (i.e., a point and itself) on the sphere is zero. ^ σφαῖρα, Henry George Liddell, Robert Scott, A Greek-English Lexicon, on Perseus. ^ a b Albert 2016, p. 54. ^ a b c d e f Chisholm, Hugh, ed. (1911). "Sphere" . Encyclopædia Britannica. Vol. 25 (11th ed.). Cambridge University Press. pp. 647–648. ^ a b c Woods 1961, p. 266. ^ Kreyszig (1972, p. 342). ^ Steinhaus 1969, p. 223. ^ "The volume of a sphere – Math Central". mathcentral.uregina.ca. Retrieved 10 June 2019. ^ a b E.J. Borowski; J.M. Borwein (1989). Collins Dictionary of Mathematics. Collins. pp. 141, 149. ISBN 978-0-00-434347-1. ^ a b c Weisstein, Eric W. "Sphere". MathWorld. ^ Steinhaus 1969, p. 221. ^ Osserman, Robert (1978). "The isoperimetric inequality". Bulletin of the American Mathematical Society. 84 (6): 1187. doi:10.1090/S0002-9904-1978-14553-4. Retrieved 14 December 2019. ^ Albert 2016, p. 60. ^ Albert 2016, p. 55. ^ Albert 2016, p. 57. ^ Woods 1961, p. 267. ^ Albert 2016, p. 58. ^ Hilbert, David; Cohn-Vossen, Stephan (1952). "Eleven properties of the sphere". Geometry and the Imagination (2nd ed.). Chelsea. pp. 215–231. ISBN 978-0-8284-1087-8. {{cite book}}: ISBN / Date incompatibility (help) ^ Weisstein, Eric W. "Spheric section". MathWorld. ^ "Loxodrome". ^ Fried, Michael N. (25 February 2019). "conic sections". Oxford Research Encyclopedia of Classics. doi:10.1093/acrefore/9780199381135.013.8161. ISBN 978-0-19-938113-5. Retrieved 4 November 2022. More significantly, Vitruvius (On Architecture, Vitr. 9.8) associated conical sundials with Dionysodorus (early 2nd century bce), and Dionysodorus, according to Eutocius of Ascalon (c. 480–540 ce), used conic sections to complete a solution for Archimedes' problem of cutting a sphere by a plane so that the ratio of the resulting volumes would be the same as a given ratio. ^ New Scientist | Technology | Roundest objects in the world created. Wikisource has the text of the 1911 Encyclopædia Britannica article "Sphere". Sphere at Wikipedia's sister projects Definitions from WiktionaryMedia from CommonsNews from WikinewsQuotations from WikiquoteTexts from WikisourceTextbooks from WikibooksResources from Wikiversity Albert, Abraham Adrian (2016) [1949], Solid Analytic Geometry, Dover, ISBN 978-0-486-81026-3. Dunham, William (1997). The Mathematical Universe: An Alphabetical Journey Through the Great Proofs, Problems and Personalities. New York: Wiley. pp. 28, 226. Bibcode:1994muaa.book.....D. ISBN 978-0-471-17661-9. Kreyszig, Erwin (1972), Advanced Engineering Mathematics (3rd ed.), New York: Wiley, ISBN 978-0-471-50728-4. Steinhaus, H. (1969), Mathematical Snapshots (Third American ed.), Oxford University Press. Woods, Frederick S. (1961) [1922], Higher Geometry / An Introduction to Advanced Methods in Analytic Geometry, Dover. John C. Polking (15 April 1999). "The Geometry of the Sphere". www.math.csi.cuny.edu. Retrieved 21 January 2022. Mathematica/Uniform Spherical Distribution Surface area of sphere proof Retrieved from " 3Mathematical object This article includes a list of general references, but it lacks sufficient corresponding inline citations. Please help to improve this article by introducing more precise citations. (June 2016) (Learn how and when to remove this message) Stereographic projection of the hypersphere's parallels (red), meridians (blue) and hypermeridians (green). Because this projection is conformal, the curves intersect each other orthogonally (in the yellow points) as in 4D. All curves are circles: the curves that intersect ⟨0,0,0,1⟩ have infinite radius (= straight line). In this picture, the whole 3D space maps the surface of the hypersphere, whereas in the next picture the 3D space contained the shadow of the bulk hypersphere. Direct projection of 3-sphere into 3D space and covered with surface grid, showing structure as stack of 3D spheres (2-spheres) In mathematics, a hypersphere or 3-sphere is a 4-dimensional analogue of a sphere, and is the 3-dimensional n-sphere. In 4-dimensional Euclidean space, it is the set of points equidistant from a fixed central point. The interior of a 3-sphere is a 4-ball. It is called a 3-sphere because topologically, the surface itself is 3-dimensional, even though it is curved into the 4th dimension. For example, when traveling on a 3-sphere, you can go north and south, east and west, or along a 3rd set of cardinal directions. This means that a 3-sphere is an example of a 3-manifold. In coordinates, a 3-sphere with center (C0, C1, C2, C3) and radius r is the set of all points (x0, x1, x2, x3) in real, 4-dimensional space (R4) such that ∑ i = 0 3 (x i − C i) 2 = (x 0 − C 0) 2 + (x 1 − C 1) 2 + (x 2 − C 2) 2 + (x 3 − C 3) 2 = r 2 . {\displaystyle \sum _{i=0}^{3}(x_{i}-C_{i})^{2}=(x_{0}-C_{0})^{2}+(x_{1}-C_{1})^{2}+(x_{2}-C_{2})^{2}+(x_{3}-C_{3})^{2}=r^{2}.} The 3-sphere centered at the origin with radius 1 is called the unit 3-sphere and is usually denoted S3: S 3 = { (x 0 , x 1 , x 2 , x 3) ∈ R 4 : x 0 2 + x 1 2 + x 2 2 + x 3 2 = 1 } . {\displaystyle S^{3}=\left\{(x_{0},x_{1},x_{2},x_{3})\in \mathbb {R} ^{4}:x_{0}^{2}+x_{1}^{2}+x_{2}^{2}+x_{3}^{2}=1\right\}.} It is often convenient to regard R4 as the space with 2 complex dimensions (C2) or the quaternions (H). The unit 3-sphere is then given by S 3 = { (z 1 , z 2) ∈ C 2 : | z 1 | 2 + | z 2 | 2 = 1 } {\displaystyle S^{3}=\left\{(z_{1},z_{2})\in \mathbb {C} ^{2}:|z_{1}|^{2}+|z_{2}|^{2}=1\right\}} or S 3 = { q ∈ H : ‖ q ‖ = 1 } . {\displaystyle S^{3}=\left\{q\in \mathbb {H} :\|q\|=1\right\}.} This description as the quaternions of norm one identifies the 3-sphere with the versors in the quaternion division ring. Just as the unit circle is important for planar polar coordinates, so the 3-sphere is important in the polar view of 4-space involved in quaternion multiplication. See polar decomposition of a quaternion for details of this development of the three-sphere. This view of the 3-sphere is the basis for the study of elliptic space as developed by Georges Lemaître.[1] The 3-dimensional surface volume of a 3-sphere of radius r is S V = 2 π 2 r 3 {\displaystyle SV=2\pi ^{2}r^{3}\,} while the 4-dimensional hypervolume (the content of the 4-dimensional region, or ball, bounded by the 3-sphere) is H = 1 2 π 2 r 4 . {\displaystyle H={\frac {1}{2}}\pi ^{2}r^{4}.} Every non-empty intersection of a 3-sphere with a three-dimensional hyperplane is a 2-sphere (unless the hyperplane is tangent to the 3-sphere, in which case the intersection is a single point). As a 3-sphere moves through a given three-dimensional hyperplane, the intersection starts out as a point, then becomes a growing 2-sphere that reaches its maximal size when the hyperplane cuts right through the "equator" of the 3-sphere. Then the 2-sphere shrinks again down to a single point as the 3-sphere leaves the hyperplane. In a given three-dimensional hyperplane, a 3-sphere can rotate about an "equatorial plane" (analogous to a 2-sphere rotating about a central axis), in which case it appears to be a 2-sphere whose size is constant. A 3-sphere is a compact, connected, 3-dimensional manifold without boundary. It is also simply connected. What this means, in the broad sense, is that any loop, or circular path, on the 3-sphere can be continuously shrunk to a point without leaving the 3-sphere. The Poincaré conjecture, proved in 2003 by Grigori Perelman, provides that the 3-sphere is the only three-dimensional manifold (up to homeomorphism) with these properties. The 3-sphere is homeomorphic to the one-point compactification of R3. In general, any topological space that is homeomorphic to the 3-sphere is called a topological 3-sphere. The homology groups of the 3-sphere are as follows: H0(S3, Z) and H3(S3, Z) are both infinite cyclic, while Hi(S3, Z) = {} for all other indices i. Any topological space with these homology groups is known as a homology 3-sphere. Initially Poincaré conjectured that all homology 3-spheres are homeomorphic to S3, but then he himself constructed a non-homeomorphic one, now known as the Poincaré homology sphere. Infinitely many homology spheres are now known to exist. For example, a Dehn filling with slope 1/n on any knot in the 3-sphere gives a homology sphere; typically these are not homeomorphic to the 3-sphere. As to the homotopy groups, we have π1(S3) = π2(S3) = {} and π3(S3) is infinite cyclic. The higher-homotopy groups (k ≥ 4) are all finite abelian but otherwise follow no discernible pattern. For more discussion see homotopy groups of spheres. Homotopy groups of S3 k 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 πk(S3) 0 0 0 Z Z2 Z2 Z12 Z2 Z2 Z3 Z15 Z2 Z2⊕Z2 Z12⊕Z2 Z84⊕Z2⊕Z2 Z2⊕Z2 Z6 The 3-sphere is naturally a smooth manifold, in fact, a closed embedded submanifold of R4. The Euclidean metric on R4 induces a metric on the 3-sphere giving it the structure of a Riemannian manifold. As with all spheres, the 3-sphere has constant positive sectional curvature equal to 1/r2 where r is the radius. Much of the interesting geometry of the 3-sphere stems from the fact that the 3-sphere has a natural Lie group structure given by quaternion multiplication (see the section below on group structure). The only other spheres with such a structure are the 0-sphere and the 1-sphere (see circle group). Unlike the 2-sphere, the 3-sphere admits nonvanishing vector fields (sections of its tangent bundle). One can even find three linearly independent and nonvanishing vector fields. These may be taken to be any left-invariant vector fields forming a basis for the Lie algebra of the 3-sphere. This implies that the 3-sphere is parallelizable. It follows that the tangent bundle of the 3-sphere is trivial. For a general discussion of the number of linear independent vector fields on a n-sphere, see the article vector fields on spheres. There is an interesting action of the circle group T on S3 giving the 3-sphere the structure of a principal circle bundle known as the Hopf bundle. If one thinks of S3 as a subset of C2, the action is given by (z 1 , z 2) · λ = (z 1 λ , z 2 λ) ∀ λ ∈ T {\displaystyle (z_{1},z_{2})\cdot \lambda =(z_{1}\lambda ,z_{2}\lambda)\quad \forall \lambda \in \mathbb {T} } . The orbit space of this action is homeomorphic to the two-sphere S2. Since S3 is not homeomorphic to S2 × S1, the Hopf bundle is nontrivial. There are several well-known constructions of the three-sphere. Here we describe gluing a pair of three-balls and then the one-point compactification. A 3-sphere can be constructed topologically by "gluing" together the boundaries of a pair of 3-balls. The boundary of a 3-ball is a 2-sphere, and these two 2-spheres are to be identified. That is, imagine a pair of 3-balls of the same size, then superpose them so that their 2-spherical boundaries match, and let matching pairs of points on the pair of 2-spheres be identically equivalent to each other. In analogy with the case of the 2-sphere (see below), the gluing surface is called an equatorial sphere. Note that the interiors of the 3-balls are not glued to each other. One way to think of the fourth dimension is as a continuous real-valued function of the 3-dimensional coordinates of the 3-ball, perhaps considered to be "temperature". We take the "temperature" to be zero along the gluing 2-sphere and let one of the 3-balls be "hot" and let the other 3-ball be "cold". The "hot" 3-ball could be thought of as the "upper hemisphere" and the "cold" 3-ball could be thought of as the "lower hemisphere". The temperature is highest/lowest at the centers of the two 3-balls. This construction is analogous to a construction of a 2-sphere, performed by gluing the boundaries of a pair of disks. A disk is a 2-ball, and the boundary of a disk is a circle (a 1-sphere). Let a pair of disks be of the same diameter. Superpose them and glue corresponding points on their boundaries. Again one may think of the third dimension as temperature. Likewise, we may inflate the 2-sphere, moving the pair of disks to become the northern and southern hemispheres. After removing a single point from the 2-sphere, what remains is homeomorphic to the Euclidean plane. In the same way, removing a single point from the 3-sphere yields three-dimensional space. An extremely useful way to see this is via stereographic projection. We first describe the lower-dimensional version. Rest the south pole of a unit 2-sphere on the xy-plane in three-space. We map a point P of the sphere (minus the north pole N) to the plane by sending P to the intersection of the line NP with the plane. Stereographic projection of a 3-sphere (again removing the north pole) maps to three-space in the same manner. (Notice that, since stereographic projection is conformal, round spheres are sent to round spheres or to planes.) A somewhat different way to think of the one-point compactification is via the exponential map. Returning to our picture of the unit two-sphere sitting on the Euclidean plane: Consider a geodesic in the plane, based at the origin, and map this to a geodesic in the two-sphere of the same length, based at the south pole. Under this map all points of the circle of radius π are sent to the north pole. Since the open unit disk is homeomorphic to the Euclidean plane, this is again a one-point compactification. The exponential map for 3-sphere is similarly constructed; it may also be discussed using the fact that the 3-sphere is the Lie group of unit quaternions. The four Euclidean coordinates for S3 are redundant since they are subject to the condition that x02 + x12 + x22 + x32 = 1. As a 3-dimensional manifold one should be able to parameterize S3 by three coordinates, just as one can parameterize the 2-sphere using two coordinates (such as latitude and longitude). Due to the nontrivial topology of S3 it is impossible to find a single set of coordinates that cover the entire space. Just as on the 2-sphere, one must use at least two coordinate charts. Some different choices of coordinates are given below. It is convenient to have some sort of hyperspherical coordinates on S3 in analogy to the usual spherical coordinates on S2. One such choice — by no means unique — is to use (ψ, θ, φ), where x 0 = r cos ψ x 1 = r sin ψ cos θ x 2 = r sin ψ sin θ cos φ x 3 = r sin ψ sin θ sin φ {\displaystyle {\begin{aligned}x_{0}&=r\cos \psi \\x_{1}&=r\sin \psi \cos \theta \\x_{2}&=r\sin \psi \sin \theta \cos \varphi \\x_{3}&=r\sin \psi \sin \theta \sin \varphi \end{aligned}}} where ψ and θ run over the range 0 to π, and φ runs over 0 to 2π. Note that, for any fixed value of ψ, θ and φ parameterize a 2-sphere of radius r sin ψ {\displaystyle r\sin \psi } , except for the degenerate cases, when ψ equals 0 or π, in which case they describe a point. The round metric on the 3-sphere in these coordinates is given by[2] d s 2 = r 2 [d ψ 2 + sin 2 ψ (d θ 2 + sin 2 θ d φ 2)] {\displaystyle ds^{2}=r^{2}\left[d\psi ^{2}+\sin ^{2}\psi \left(d\theta ^{2}+\sin ^{2}\theta \,d\varphi ^{2}\right)\right]} and the volume form by d V = r 3 (sin 2 ψ sin θ) d ψ ∧ d θ ∧ d φ . {\displaystyle dV=r^{3}\left(\sin ^{2}\psi \,\sin \theta \right)\,d\psi \wedge d\theta \wedge d\varphi .} These coordinates have an elegant description in terms of quaternions. Any unit quaternion q can be written as a versor: q = e τ ψ = cos ψ +τ sin ψ {\displaystyle q=e^{\tau \psi }=\cos \psi +\tau \sin \psi } where τ is a unit imaginary quaternion; that is, a quaternion that satisfies τ2 = −1. This is the quaternionic analogue of Euler's formula. Now the unit imaginary quaternions all lie on the unit 2-sphere in Im H so any such τ can be written: τ = (cos θ) i + (sin θ cos φ) j + (sin θ sin φ) k {\displaystyle \tau =(\cos \theta)i+(\sin \theta \cos \varphi)j+(\sin \theta \sin \varphi)k} With τ in this form, the unit quaternion q is given by q = e τ ψ = x 0 + x 1 i + x 2 j + x 3 k {\displaystyle q=e^{\tau \psi }=x_{0}+x_{1}i+x_{2}j+x_{3}k} where x0,1,2,3 are as above. When q is used to describe spatial rotations (cf. quaternions and spatial rotations), it describes a rotation about τ through an angle of 2ψ. The Hopf fibration can be visualized using a stereographic projection of S3 to R3 and then compressing R3 to a ball. This image shows points on S2 and their corresponding fibers with the same color. For unit radius another choice of hyperspherical coordinates, (η, ξ1, ξ2), makes use of the embedding of S3 in C2. In complex coordinates (z1, z2) ∈ C2 we write z 1 = e i ξ 1 sin η z 2 = e i ξ 2 cos η . {\displaystyle {\begin{aligned}z_{1}&=e^{i\,\xi _{1}}\sin \eta \\z_{2}&=e^{i\,\xi _{2}}\cos \eta .\end{aligned}}} This could also be expressed in R4 as x 0 = cos ξ 1 sin η x 1 = sin ξ 1 sin η x 2 = cos ξ 2 cos η x 3 = sin ξ 2 cos η . {\displaystyle {\begin{aligned}x_{0}&=\cos \xi _{1}\sin \eta \\x_{1}&=\sin \xi _{1}\sin \eta \\x_{2}&=\cos \xi _{2}\cos \eta \\x_{3}&=\sin \xi _{2}\cos \eta .\end{aligned}}} Here η runs over the range 0 to π/2, and ξ1 and ξ2 can take any values between 0 and 2π. These coordinates are useful in the description of the 3-sphere as the Hopf bundle S 1 → S 3 → S 2 . {\displaystyle S^{1}\to S^{3}\to S^{2}.\,} A diagram depicting the poloidal (ξ1) direction, represented by the red arrow, and the toroidal (ξ2) direction, represented by the blue arrow, although the terms poloidal and toroidal are arbitrary in this flat torus case. For any fixed value of η between 0 and π/2, the coordinates (ξ1, ξ2) parameterize a 2-dimensional torus. Rings of constant ξ1 and ξ2 above form simple orthogonal grids on the torii. See image to right. In the degenerate cases, when η equals 0 or π/2, these coordinates describe a circle. The round metric on the 3-sphere in these coordinates is given by d s 2 = d η 2 + sin 2 η d ξ 1 2 + cos 2 η d ξ 2 2 {\displaystyle ds^{2}=d\eta ^{2}+\sin ^{2}\eta \,d\xi _{1}^{2}+\cos ^{2}\eta \,d\xi _{2}^{2}} and the volume form by d V = sin η cos η d η ∧ d ξ 1 ∧ d ξ 2 . {\displaystyle dV=\sin \eta \cos \eta \,d\eta \wedge d\xi _{1}\wedge d\xi _{2}.} To get the interlocking circles of the Hopf fibration, make a simple substitution in the equations above[3] z 1 = e i (ξ 1 + ξ 2) sin η z 2 = e i (ξ 2 − ξ 1) cos η . {\displaystyle {\begin{aligned}z_{1}&=e^{i\,(\xi _{1}+\xi _{2})}\sin \eta \\z_{2}&=e^{i\,(\xi _{2}-\xi _{1})}\cos \eta .\end{aligned}}} In this case η, and ξ1 specify which circle, and ξ2 specifies the position along each circle. One round trip (0 to 2π) of ξ1 or ξ2 equates to a round trip of the torus in the 2 respective directions. Another convenient set of coordinates can be obtained via stereographic projection of S3 from a pole onto the corresponding equatorial R3 hyperplane. For example, if we project from the point (−1, 0, 0, 0) we can write a point p in S3 as p = (1 − ‖ u ‖ 2 1 + ‖ u ‖ 2 , 2 u 1 + ‖ u ‖ 2) = 1 + u 1 − u {\displaystyle p=\left({\frac {1-\|u\|^{2}}{1+\|u\|^{2}}},{\frac {2\mathbf {u} }{1+\|u\|^{2}}}\right)={\frac {1+\mathbf {u} }{1-\mathbf {u} }}} where u = (u1, u2, u3) is a vector in R3 and ‖u‖2 = u12 + u22 + u32. In the second equality above, we have identified p with a unit quaternion and u = u1i + u2j + u3k with a pure quaternion. (Note that the numerator and denominator commute here even though quaternionic multiplication is generally noncommutative). The inverse of this map takes p = (x0, x1, x2, x3) in S3 to u = 1 1 + x 0 (x 1 , x 2 , x 3) . {\displaystyle \mathbf {u} ={\frac {1}{1+x_{0}}}\left(x_{1},x_{2},x_{3}\right).} We could just as well have projected from the point (1, 0, 0, 0), in which case the point p is given by p = (− 1 + ‖ v ‖ 2 1 + ‖ v ‖ 2 , 2 v 1 + ‖ v ‖ 2) = − 1 + v 1 + v {\displaystyle p=\left({\frac {-1+\|v\|^{2}}{1+\|v\|^{2}}},{\frac {2\mathbf {v} }{1+\|v\|^{2}}}\right)={\frac {-1+\mathbf {v} }{1+\mathbf {v} }}} where v = (v1, v2, v3) is another vector in R3. The inverse of this map takes p to v = 1 1 − x 0 (x 1 , x 2 , x 3) . {\displaystyle \mathbf {v} ={\frac {1}{1-x_{0}}}\left(x_{1},x_{2},x_{3}\right).} Note that the u coordinates are defined everywhere but (−1, 0, 0, 0) and the v coordinates everywhere but (1, 0, 0, 0). This defines an atlas on S3 consisting of two coordinate charts or "patches", which together cover all of S3. Note that the transition function between these two charts on their overlap is given by v = 1 ‖ u ‖ 2 u {\displaystyle \mathbf {v} ={\frac {1}{\|u\|^{2}}}\mathbf {u} } and vice versa. When considered as the set of unit quaternions, S3 inherits an important structure, namely that of quaternionic multiplication. Because the set of unit quaternions is closed under multiplication, S3 takes on the structure of a group. Moreover, since quaternionic multiplication is smooth, S3 can be regarded as a real Lie group. It is a nonabelian, compact Lie group of dimension 3. When thought of as a Lie group, S3 is often denoted Sp(1) or U(1, H). It turns out that the only spheres that admit a Lie group structure are S1, thought of as the set of unit complex numbers, and S3, the set of unit quaternions (The degenerate case S0 which consists of the real numbers 1 and −1 is also a Lie group, albeit a 0-dimensional one). One might think that S7, the set of unit octonions, would form a Lie group, but this fails since octonion multiplication is nonassociative. The octonionic structure does give S7 one important property: parallelizability. It turns out that the only spheres that are parallelizable are S1, S3, and S7. By using a matrix representation of the quaternions, H, one obtains a matrix representation of S3. One convenient choice is given by the Pauli matrices: x 1 + x 2 i + x 3 j + x 4 k ↦ (x 1 + i x 2 x 3 + i x 4 − x 3 + i x 4 x 1 − i x 2) . {\displaystyle x_{1}+x_{2}i+x_{3}j+x_{4}k\mapsto {\begin{pmatrix}\;\;\,x_{1}+ix_{2}&x_{3}+ix_{4}\\-x_{3}+ix_{4}&x_{1}-ix_{2}\end{pmatrix}}.} This map gives an injective algebra homomorphism from H to the set of 2 × 2 complex matrices. It has the property that the absolute value of a quaternion q is equal to the square root of the determinant of the matrix image of q. The set of unit quaternions is then given by matrices of the above form with unit determinant. This matrix subgroup is precisely the special unitary group SU(2). Thus, S3 as a Lie group is isomorphic to SU(2). Using our Hopf coordinates (η, ξ1, ξ2) we can then write any element of SU(2) in the form (e i ξ 1 sin η e i ξ 2 cos η − e − i ξ 2 cos η e − i ξ 1 sin η) . {\displaystyle {\begin{pmatrix}e^{i\,\xi _{1}}\sin \eta &e^{i\,\xi _{2}}\cos \eta \\-e^{-i\,\xi _{2}}\cos \eta &e^{-i\,\xi _{1}}\sin \eta \end{pmatrix}}.} Another way to state this result is if we express the matrix representation of an element of SU(2) as an exponential of a linear combination of the Pauli matrices. It is seen that an arbitrary element U ∈ SU(2) can be written as U = exp (∑ i = 1 3 α i J i) . {\displaystyle U=\exp \left(\sum _{i=1}^{3}\alpha _{i}J_{i}\right).} [4] The condition that the determinant of U is +1 implies that the coefficients α1 are constrained to lie on a 3-sphere. In Edwin Abbott Abbott's Flatland, published in 1884, and in Sphereland, a 1965 sequel to Flatland by Dionys Burger, the 3-sphere is referred to as an oversphere, and a 4-sphere is referred to as a hypersphere. Writing in the American Journal of Physics,[5] Mark A. Peterson describes three different ways of visualizing 3-spheres and points out language in The Divine Comedy that suggests Dante viewed the Universe in the same way; Carlo Rovelli supports the same idea.[6] In Art Meets Mathematics in the Fourth Dimension,[7] Stephen L. Lipscomb