Glossary

Aleph: the first letter of the Hebrew alphabet, $\aleph$ is the only symbol to be used in mathematics in exclusively one way: to denote cardinal numbers. See cardinality.

Algorithm: a procedure which takes an input and computes a corresponding output in a finite number of steps. There are various ways of making precise the idea of a 'procedure', generally assumed to be equivalent under the *Church-Turing thesis* (more detail).

Examples: Cook's Theorem

Asymptotic: meaning 'approaching arbitrarily close' in the sense that the ratio of two quantities tends to one. Thus we require that $A/B \to 1$ even though, meanwhile, it may happen that $|A - B| \to \infty$.

Examples: Stirling's Approximation, Prime Number Theorem

Bijection: a function that is injective (no two things are mapped to the same codomain member) and surjective (every member of the codomain is assigned a member of the domain). Bijective functions are invertible.

Cardinality: the number of elements in a set, denoted for set X by $|X|$, occasionally by $\#X$. For finite sets this presents no difficulties; infinite sets can still have their sizes compared by means of constructing one-to-one mappings (see under functions).

The size of the integers $\mathbb{Z}$ is denoted by $\aleph_0$ and, for any $i \geq 0$, $\aleph_{i+1}$ is used to denote the next largest cardinal after $\aleph_i$. This certainly exists since, if X is an infinite set, then a set of greater cardinality is constructed as the powerset of X. For $\mathbb{Z}$ we obtain $2^{\mathbb{Z}} = \mathbb{R}$. However, it is not known if $|\mathbb{R}| = \aleph_1$, this being the assertion of the *continuum hypothesis*.

Examples: Cantor's Uncountability Theorem, Cantor's Theorem

Chromatic number: for a graph G, denoted by $\chi(G)$: the smallest number of colours needed so that the vertices of G can be coloured with no edge joining vertices of the same colour.

Examples: The Four Colour Theorem

Closed: an operation upon members of a mathematical structure (e.g. group or ring is closed if the result of the operation is again a member of the same structure. E.g. adding two integers gives again an integer; division in the set of integers, on the other hand, is not closed as it may take us outside the integers.

Complex: involving complex numbers, i.e. those which as well as having a real number part, have an imaginary part: a multiple of $i = \sqrt{-1}$.

Examples: The Fundamental Theorem of Algebra, The First Isomorphism Theorem

Complexity: some measure of the amount of information contained in, or required by, a mathematical object or problem. *Computational complexity* is a measure of how inherently hard a problem is to solve computationally. Algorithmic complexity is a measure of how many time steps a given algorithm will take to solve a given problem in the average/worst case. *Kolmogorov complexity* is a measure of the amount of resource needed to specify a mathematical object. There are many more.

Examples: Cook's Theorem

Constructive: of proofs: meaning an object is proved to exist by showing how it may be constructed. For instance, a constructive proof that the rationals are countable requires an explicit method for putting them into one-to-one correspondence with the positive integers (as here, for example). Bob Lockhart has told me of a wonderful example of non-constructivism: "Are there irrational numbers which, raised to the power of a suitable irrational number, give rational numbers? Try $\sqrt{2}^{\sqrt{2}}$. If this is rational, we are done; if it is irrational, then raising it to the power $\sqrt{2}$ gives $\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \sqrt{2}^{\sqrt{2}\times\sqrt{2}} = \sqrt{2}^{2} = 2$, a rational and we are done. Brouwer would say this is not a proof because we have not exhibited an irrational number with the required properties."

Examples: Lagrange's Four Squares Theorem, Brouwer's Fixed Point Theorem

Continuous: taking values from an uncountable set, usually some subset of the real numbers $\mathbb{R}$.

Convergence: a sequence $x_1, x_2, \ldots$ of members of an infinite subset X of a set Y is said to converge to a limit $y \in Y$ if, however small a distance we specify, our x_i's eventually all lie at least that close to y. We can then write $\lim_{i\to\infty} x_i = y$. This is made precise in a typical 'epsilon-type' formulation: for any $\varepsilon > 0$ there is some N such that, for all $i \geq N$, $|x_i - y| < \varepsilon$. (Consequently epsilon has entered the mathematician's general vocabulary: "I got promoted but they only increased my salary by epsilon." Paul Erdös took this one further and referred to all children affectionately as epsilons.)

Examples: Stirling's Approximation, The Prime Number Theorem

Coset: if G is a group and H is a subgroup of G, then the *right* cosets of H in G are the collection the sets Hg. For $g_1 \neq g_2 \in G$ it may or may not happen that $Hg_1 = Hg_2$. However, these sets will either be identical or disjoint from each other (the cosets partition G). *Left* cosets are defined analogously (the sets gH).

Examples: Lagrange's Theorem, First Isomorphism Theorem

Countable: a set of objects is called *countable* if there is a way of listing its members as: first member, second member, third member, and so on. If the set has infinitely many members, for example the set of positive integers, $\{1, 2, 3, \ldots\}$, then it may be referred to as *countably infinite* to distinguish it from set such as the real numbers $\mathbb{R}$ which are *uncountable*.

Examples: Cantor's Uncountability Theorem, Cantor's Theorem

Degrees: measurement of angle by dividing the circle into 360 equal segments. Although degrees are more common in every day use, radians are, mathematically, much to be preferred (see why). Just as between Celsius and Fahrenheit there is a simple conversion: Radians = Degrees $\times \pi/180$.

Examples: Euler's Identity, Pythagoras' Theorem

Determinant: a function which assigns to a square matrix a single value of the same numerical type as the matrix entries (real number, complex number, etc). Viewing an $n \times n$ matrix as a linear transformation in n-dimensional space, the determinant can be interpreted geometrically as being the factor by which the volume of an n-dimensional solid is increased or decreased under the transformation.

Examples: Matrix Tree Theorem

Discrete: taking values from a countable set, e.g. the integers $\mathbb{Z}$ or the rationals $\mathbb{Q}$.

Examples: The Central Limit Theorem

Distribution: a function which assigns to each member of its domain a frequency (*frequency distribution*) or probability (*probability distribution*) of occurrence. In the case of a probability distribution, the members of the domain are the *outcomes* of some *trial*, which is specified in some precise way and which simultaneously gives rise to the distribution function.

More accurately, for continuous distributions, probabilities are assigned to intervals in the domain: the question “will an angle be measured to lie in the interval $[-\pi/2, \pi/2]$?” makes sense but the question “will it be measured to be *exactly* π?” does not.

Examples: The Central Limit Theorem, Benford’s Law

Elementary: not as in ‘Elementary, my dear Watson’ but as in ‘elementary proof of the prime number theorem’, meaning not requiring sophisticated mathematical machinery imported from other branches of mathematics (complex analysis, in the case of the prime number theorem). Elementary mathematical results may be far from trivial and indeed the difficult and ingenious work of elementary number theorists merits its own subject code in the American Mathematical Society’s Subject Classification.

Function: an assignment of the members in some set (the *domain*) to the members of another set (the *range* or *codomain*), obeying the rule that no member of the domain gets assigned to more than one member of the codomain (the assignment is unambiguous). If X is the domain and Y the codomain of a function f, then it is common to represent the assignment in the form

$$f : X \to Y.$$

Then the subset of Y assigned members of X is called the *image* of f, denoted $\mathrm{Im}(f)$. If the assignment targets everything in the codomain ($\mathrm{Im}(f) = Y$) then the function is called surjective or *onto*; if nothing in the codomain is assigned more than one member of the domain the function is injective or *one-to-one*.

Examples: First Isomorphism Theorem, Fundamental Theorem of Arithmetic

Fixed point: a member of the domain of a function f having the property that it is left unchanged by f. Thus x, necessarily lying in both the domain and codomain of f, satisfies $f(x) = x$.

Examples: Brouwer’s Fixed Point Theorem

Graph: 1. a set of objects called *vertices* (or *nodes*) and a set of *edges*, each joining two vertices. Variations include: *directed graphs*, where the edges (sometimes called *arcs*) have a direction; edges called *loops* which join a vertex to itself; *multiple edges* in which more than one edge may connect any pair of vertices.

Examples: Euler's Formula, Four Colour Theorem

2. a representation of a functional relationship between two sets. The items of one set are laid out along a *horizontal axis* line and the items of the other along a *vertical axis*. If the function is f and it maps item x to item y, say, then $f(x) = y$ is represented as a dot above x and to the right of y. In the classic case where the two sets are the real numbers $\mathbb{R}$, the uncountable sequence of dots is replaced by a continuous line.

   Examples: Merton College Theorem, Prime Number Theorem

Group: a set (whose members are called the group *elements*) together with a rule (usually called *(group) multiplication*) for combining any two elements to get a third. It is required of this multiplication that

1. it is *associative*: multiplying this third element by yet another gives the same end result as if the second and third were multiplied first and then multiplied by the first; in other words we can ignore brackets, since $(g_1 \times g_2) \times g_3 = g_1 \times (g_2 \times g_3)$;
2. it has among the group elements a unique *identity element* whose effect, on multiplying, is to leave unchanged any element of the group; and
3. any element has an *inverse* which multiplies with it to produce the identity.

The set of integers, with addition taking the role of multiplication, and with zero acting as the identity is a group; the set of integers using the usual multiplication is *not* since although integer multiplication is associative and has an identity (namely, 1) no element has an inverse: to get 1 from 3 we must multiply by 1/3, which is not an integer.

Examples: Orbit Counting Lemma, Lagrange's Theorem, First Isomorphism Theorem

Hard problem: somewhat dependent on the context but in mathematics this tends to mean 'not solved by any fast algorithm' which in turn has technical connotations to do with computational complexity. Thus we say that the security of the RSA encryption system relies on factorisation being a hard problem with the technical meaning that no polynomial-time algorithm is known to solve the factorisation problem.

Examples: Cook's Theorem

Homomorphism: literally 'same form'; a function which maps one mathematical object to another in such a way that structure is preserved. A homomorphism between two graphs (1), for example, consists of a mapping of the vertices of one to those of another in such a way that edges are mapped to edges: if h is the function and u and v are two vertices such that uv is an edge then $h(u)h(v)$ is again an edge in the target graph.

For groups it is multiplication that must be preserved: $h(a \times b)$ must be equal to $h(a) \times h(b)$.

Examples: First Isomorphism Theorem

Imaginary: being a multiple of $i = \sqrt{-1}$. Complex numbers of the form $a + ib$ are sometimes referred to as being 'purely imaginary' when $a = 0$.

Indeterminate: an unknown: an example of the famous x or y of algebra. There is a subtle difference between an indeterminate and a variable, which is largely a matter of intent: a 'variable' is something you want to give a value to; an 'indeterminate' is something whose manipulation is, in itself, meant to provide information.

Interval: a continuous portion of the real number line. If the interval runs from a to b then it may be:

Open: meaning all numbers from, but not including, a up to, but not including, b; denoted (a, b);

Closed: meaning all numbers from a to b inclusive;

Half open/half closes: either a or b is excluded: $(a, b]$ or $[a, b)$.

Either a or b can take the value ∞ (in which case the interval is open at that end).

More generally the notion of interval may be applied to any totally ordered set and, in higher dimensions, to closed or open *balls*.

Examples: Brouwer's Fixed Point Theorem

Injection: a function is an injection, or is *one-to-one*, if not two members of its domain are mapped to the same member of its codomain. The function $x \to x^2$ is an injection if the domain is the positive integers but not if the domain is the whole of $\mathbb{Z}$ since then -2 and 2, for example, both map to 4.

Inverse: of a function $f : X \to Y$, usually denoted by $f^{-1} : Y \to X$. The function which reverses the assignment of f: if f maps x to y then (and only then) f^{-1} maps y to x. The inverse function exists if and only if f is a bijection.

Invertible: a function is invertible if it has an inverse. A square matrix is invertible if it has an inverse in the sense of being a linear transformation function. This occurs precisely when the matrix has non-zero determinant.

Isomorphism: a homomorphism which preserves size as well as structure, in the technical sense that the function concerned must be a bijection.

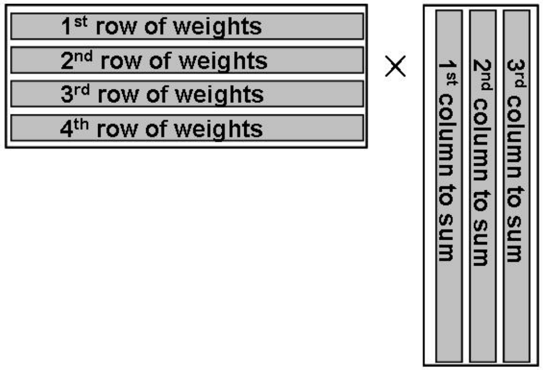
Examples: First Isomorphism Theorem, Second Isomorphism Theorem

Limit: see Convergence.

Linear transformation: a function which scales a point in n-dimensional space, the scale factor in each axis being a weighted sum of the point's coordinates. The weights can be represented using an $n \times n$ matrix. For instance if, given a particular set of axes in 2 dimensions, a transformation moves the point (x, y) to $(ax + by, cx + dy)$ then, for these axes, the transformation can be represented by the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Matrix: a collection of $m \times n$ numbers or indeterminates arranged in a rectangular array of m *rows* and n *columns*. The matrix is said to have *size* $m \times n$ (said 'm by n'). When $m = n$ the matrix is called *square*. Matrices may be added and multiplied and, especially for square matrices, this gives us a new kind of arithmetic in which many of the properties of integer or real-number arithmetic still hold.

Geometrically, an $n \times n$ matrix can be thought of as a linear transformation and *matrix multiplication* is defined accordingly using weighted sums as illustrated on the right, where $4 \times 3 = 12$ weighted sums are produced, generating the entries of a 4×3 matrix. Notice that multiplying an $m \times n$ matrix with an $r \times s$ one requires that $n = r$, since the number of weights must match the lengths of the columns.



Examples: Matrix Tree Theorem

Non-constructive: see Constructive

Partial order: a *binary relation* or ordered pairing placed upon a set of objects so as to satisfy:

Reflexivity: every object is paired with itself;

Antisymmetry: if object x is paired with object y (so (x, y) is in the binary relation) then y is *not* paired with x ((y, x) is not in the relation);

Transitivity: if (x, y) is in the relation and (y, z) is in the relation then (x, z) is in the relation.

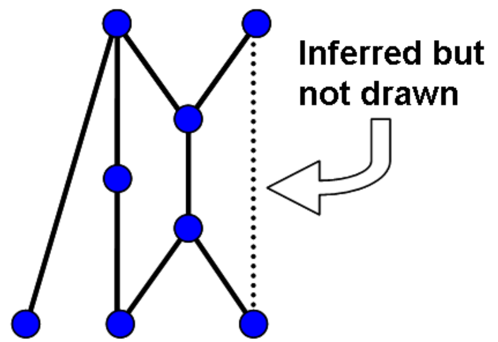
A set with such an ordering placed upon is called a partially ordered set or *poset*.

The order relation is often represented with $\leq$ in analogy with the familiar partial ordering 'less than or equal' placed upon the real numbers. The presence of (x, y) in the order relation is then more simply written $x \leq y$. The real numbers are *totally ordered* by $\leq$ (with its usual meaning) since any two numbers x and y satisfy $x \leq y$ or $y \leq x$. In general this need not be the case: two objects x and y for which neither or the pairs (x, y) and (y, x) appear in the relation are called *incomparable*.

Examples: Dilworth's Theorem

Partially ordered set or poset: a set together with a partial ordering, often denoted by $\leq$, of its members.

Transitivity of the partial order means that the presence of a cycle: $x_1 \leq x_2 \leq ... \leq x_k \leq x_1$ would contradict antisymmetry. So the members of the poset can be arranged in 'ascending order', with the smallest at the bottom and the largest at the top. The *Hasse diagram* (pronounced 'Ha! Sir') of a (finite) poset is a drawing of this arrangement with only immediate $\leq$ relations being shown, so that $x \leq z$, inferred from $x \leq y$ and $y \leq z$ by transitivity, is omitted from the diagram:



Examples: Dilworth's Theorem

Planar graph: a graph (1) whose vertices and edges can be so drawn in the plane (or on the surface of a sphere) that edges meet one another only at vertices. A graph so drawn is called a *plane graph*.

Examples: Euler's Formula, Kurtowski's Theorem

Polynomial: a weighted sum of powers of an indeterminate, say x. The weight of each power is called it's *coefficient*. A *polynomial equation* is derived by setting a polynomial equal to some number (often zero) or (which amounts to the same thing) another polynomial in the same indeterminate.

Examples: Fundamental Theorem of Algebra, Euler's Identity

Polynomial-time: as referring to an algorithm: meaning that algorithm completes in an amount of time which is a polynomial function of the size of the input. Consider a room containing n people; one algorithm to check whether any of them know each other would be to bring together each pair in turn and ask "d'you two know each other?". This is a polynomial-time algorithm since there are $\binom{n}{2} = n^2/2 + n/2$ pairs. Compared to the n^2, the factors of $1/2$ and even the n are not significant and this would usually be referred to as an n^2 or quadratic algorithm.

Examples: Cook's Theorem

Poset: see Partially ordered set.

Powerset: the set built from set X by taking all subsets of X. Sometimes denoted 2^X and sometimes $\mathcal{P}(X)$.

Examples: Cantor's Theorem

Prime number: any positive integer which can only be divided exactly by itself or one, but excluding the number one itself. A fine glossary of prime number terminology can be found here. (explanation).

Examples: Euclid's Theorem, Prime Number Theorem

Probability: a real number, lying in the closed interval $[0, 1]$, specifying how likely an event is to occur. The Law of Large Numbers justifies our thinking of this as the relative frequency of that event occurring as the outcome of a trial, as the number of times the trial is repeated approaches infinity. It is sometimes more comfortable to think of a probability of, say, 0.68, as the corresponding percentage: 68% chance of occurrence. The probability of event A is denoted by $\mathrm{Prob}(A)$ or $\Pr(A)$ or (as in theorems of the day) $\mathbb{P}(A)$.

Examples: Bayes' Theorem, Benford's Law

Pseudorandom: chosen so as to simulate some probability distribution (usually the uniform distribution). Computers use sophisticated algorithms for producing, say, a number between 0 and 1, in a way which appears to be random; and there are sophisticated ways of testing how successful they are. To avoid repetition, pseudorandom number generators need a *seed* from which a number, or series of numbers, is then produced. This might, for example, be derived from the system clock.

Examples: Law of Large Numbers

Quadratic: involving numbers, variables or indeterminates raised to power of two.

Examples: Pythagoras' Theorem, Quadratic Residue Theorem

Radians: an alternative to degrees as a way of measuring angles. Just as between Celsius and Fahrenheit there is a simple conversion: Degrees = Radians $\times 180/\pi$. Radians are the accepted measurement of angle in mathematics, indeed many formulae and theorems work naturally in radians but require a messy 'fudge factor' in degrees (more detail).

Examples: Euler's Identity, Pythagoras' Theorem

Random: chosen from some set in a way which is consistent with some probability distribution defined on that set. The everyday usage of random, meaning every value is equally likely to occur, is referred to by mathematicians as *uniformly at random*, meaning drawn from the uniform distribution.

Examples: Benford's Law

Rational: being the ratio of two integers. Real numbers which are not rational are called *irrational*. The proof that not all real numbers are rational is an ancient classic that never tires of repeating:

> suppose $\sqrt{2} = a/b$ with a and b integers which are not both even (otherwise just divide each by two until this is true). Now $b\sqrt{2} = a$, so $2b^2 = a^2$, telling us that a^2 is even. This means that a is even, since a square of two odd numbers is odd. But then a^2 is actually a multiple of $2^2 = 4$, in which case $b^2 = a^2/2$ is even and b is also even. This contradicts our assumption that we could write $\sqrt{2}$ as the ratio of two integers not both even, and we conclude that $\sqrt{2}$ is irrational.

The proof that the rational numbers are countable is also very pretty (nicely presented here).

Subgraph: Any subset of the vertices of a graph (1) together with any subset of edges joining those vertices.

Subgroup: Any subset of the elements of a group such that the group multiplication is closed on this subset.

Examples: Lagrange's Theorem, Second Isomorphism Theorem

Surjection: a function is a surjection (or is *onto*) if every member of its codomain has some member of the domain mapped to it. The function $x \to x + 1$, for example, is surjective is the domain and codomain are the integers; it is not surjective if the domain and codomain are the *positive* integers since then nothing maps to 1.

Tends to: referring to some variable approaching arbitrarily close to a constant (or ∞). Typically, we might write, as $x \to \infty$, $f(x) \to 0$, ('as x tends to infinity, $f(x)$ tends to zero') meaning that, as we go arbitrarily far along the x axis, the function f takes values closer and closer to zero.

Examples: Stirling's Approximation, Prime Number Theorem

Totally ordered set: a partially ordered set for which the partial order $\leq$ is actually a total order: for any two members x and y of the set, either $x \leq y$ or $y \leq x$ (or both, in the case $x = y$).

Examples: The Well-ordering Theorem

Trivial: 1. zero or empty.

2. a term used by mathematicians to refer to a mathematical fact which is self-evidently true. It does not mean 'unimportant', however. It is a triviality a prime number greater than two must be odd; or that $\binom{n}{r} = \binom{n}{n-r}$ (since choosing r things from n is the same as *not* choosing $n - r$ of them); but these are key assumptions in countless proofs.

Uniform distribution: the distribution which assigns equal probability to every member of its domain set. When a single die is rolled the outcome should be a value drawn *uniformly at random* from the set $\{1, \ldots, 6\}$.

Variable: an unknown: an example of the famous x or y of algebra. It is usually meant to be assigned a value or to be solved for (compare to indeterminate). Thus we can talk about a complex variable: given a value from $\mathbb{C}$, the set of complex numbers; or a random variable: given a value at random from some distribution