4000/6000 Day 24 Dimension of vector spaces

Our next challenge is to give a precise meaning to statements like "Q(sqrt(2)) has dimension 2 over Q", and "Q(cuberoot(2)) has dimension 3 over Q". Intuitively, the dimension of a space V over F is the answer to the question "how many elements of V do you need to use, so that all other elements of V can be written as linear combinations of those elements, with coefficients in F ?" Since every element of Q(sqrt(2)) can be written in the form, a + bsqrt(2), where a,b, are rational numbers, the two elements {1, sqrt(2)} suffice to describe all other elements of Q(sqrt(2)) as linear combinations over Q. Since this cannot be done with fewer than two elements, the dimension of Q(sqrt(2)) over Q is 2. Making this precise and proving it, requires a bit of work, in particular some careful definitions.

Definition: A vector space over a field F is a set V in which an addition is defined which is associative and commutative, there is an additive identity, and additive inverses exist. Thus all the properties of a field hold for the operation of addition. (It may not be possible to multiply elements of V by each other.) However, it is possible to multiply elements of V by elements of the field F. This multiplication is distributive in both senses, i.e. $(a+b)v = av + bv$, if a, b, are in F and v is in V, and also $a(v+w) = av + aw$, for a in F, v,w, in V. It is also "associative" in the sense that $(ab)v = a(bv)$ for a,b, in F, v in V. We also require that $1v = v$, for all v in V. See chapter 5.1 for the complete list of axioms.

A <u>subspace</u> of a vector space is a subset which is also a vector space (with the same addition and scalar multiplication). To check whether a subset is a subspace, most of the axioms are automatic because they hold in the larger space, so it suffices to check the subset is non empty and closed under the two operations, addition of vectors, and scalar multiplication.

Our job in computing the dimension of a space, or of a subspace, is to determine the smallest possible number of vectors which suffice to describe all other vectors in the space as linear combinations. We are all familiar with the vector space R^3 over the real number field, where the three vectors (1,0,0), (0,1,0), and (0,0,1) suffice to describe all other vectors as linear combinations of these. I.e. the vector $(a,b,c) = a(1,0,0) + b(0,1,0) + c(0,0,1)$. Since this cannot be done with only two vectors, R^3 has dimension 3 over the field R.
Learning how to tell we actually have the smallest number of vectors possible to describe all others, is our main task.

Definition: If $\{v_1,....,v_r\}$ is an indexed set of elements of a vector space V, and w is any element of V, we say w "depends on" the set $\{v_1,....,v_r\}$ if there exist elements $c_1,....,c_r$ of F (called "scalars"), such that $w = c_1v_1+....+c_rv_r$, i.e. if w can be written as a linear combination of the vectors $\{v_1,....,v_r\}$ with coefficients in F.

Example: The zero vector depends on any set $\{v_1,....,v_r\}$ since we can take all the coefficients $c_1,....,c_r$ to be zero. The linear combination $0v_1+....+0v_r = 0$, in which all the coefficients are zero, is called the <u>trivial</u> linear combination.

It may be possible to write the zero vector also as a <u>non trivial</u> linear combination, for some sets

$\{v_1,....,v_r\}$. For instance if $v_1 = v_2$ (which is allowed), we can write $0 = v_1 - v_2$ where the first two coefficients are 1 and -1 (and the rest are zero).

Each of the vectors $v_1,.....,v_r$ in the set $\{v_1,....,v_r\}$, depends on the set $\{v_1,....,v_r\}$, since we can write v_1 for instance as $v_1 = 1v_1$ (and all other coefficients zero).

Exercise: Given any indexed set of vectors $S = \{v_1,....,v_r\}$ in V, the subset of V consisting of those vectors which depend on the set S, forms a subspace of V containing S. It is the smallest subspace of V containing S. We denote this subspace "span($\{v_1,....,v_r\}$)", or span(S) and call it the subspace "spanned" by the set S.

Definition: An indexed set of vectors $S = \{v_1,....,v_r\}$ in V "spans" V, if every vector in V depends on S.

Now we want to characterize when a spanning set is "minimal", i.e. as small as possible. There are two sense in which one can mean that a spanning set S is "as small as possible". One could mean that no proper subset of S still spans, or one could mean there is no other spanning set with fewer elements. Fortunately these two meanings turn out to be equivalent when taking our coefficients from a field. This is one of the nice features of vector spaces.

Notice however that when taking linear combinations of integers this would not be true. I.e. every integer is a linear combination of the integers 3 and 5 (since they are relatively prime you can get 1, and then you can get any integer), but not of either one of them alone, so this "spanning" set is minimal in the sense that no proper subset still suffices to express every integer. But it is not minimal in the second sense, because there is another set with fewer elements, namely the set $\{1\}$, which also suffices to express every integer as a linear combination.

The next definition is extremely important. Experience shows it is difficult to understand and memorize accurately, and it deserves lots of practice.

Definition: An indexed set $S = \{v_1,....,v_r\}$ of vectors is <u>independent</u> (over F) if the only linear combination that equals the zero vector, is the trivial one with all coefficients equal to zero, i.e. if $c_1v_1+....+c_rv_r = 0$, implies that all $c_i = 0$.
A set which is not independent is called <u>dependent</u>. Thus an indexed set $S = \{v_1,....,v_r\}$ is dependent if there exist scalars $c_1,....,c_r$ which are not all zero such that $c_1v_1+....+c_rv_r = 0$. I.e. S is dependent if the zero vector can be expressed as a <u>non</u> <u>trivial</u> linear combination of the vectors $\{v_1,....,v_r\}$.

(Notice the zero on the right of the equation $c_1v_1+....+c_rv_r = 0$ is the zero vector in V, while the zero in the equation $c_i = 0$ is the zero scalar. We are writing two different objects with the same symbol, and I hope it does not cause a problem. In our main application, where our vector space over F is a field E containing F, it will not be a problem since then the two objects are actually the same. I.e. the zero element of the field E will equal the zero element of the subfield F.)

Example: The set $\{(1,0),(0,1)\}$ in R^2 is independent over R since if $a(1,0)+b(0,1) = (0,0)$, then

$(a,0)+(0,b) = (a,b) = (0,0)$, so both $a = 0$ and $b = 0$. The set $\{(1,0), (0,1), (1,2)\}$ is dependent, since $(1,0) + 2(0,1) - (0,2) = (0,0)$. The indexed set $\{v_1 = (1,0), v_2 = (1,0), v_3 = (0,1)\}$ is also dependent, since the repetition allows us to write $(0,0) = v_1 - v_2 + 0v_3$, with coefficients 1,-1,0.

The next property is special to vector spaces, e.g. it would not be true for integer linear combinations.

Lemma: (1) A set $S = \{v_1,....,v_r\}$ is independent, if and only if none of the vectors in S can be written as a linear combination of the others.
(2) If S is a spanning set for V, then S is independent if and only if no proper subset of S spans V. (Thus an independent spanning set is minimal.)
Proof: (1) We will show the contrapositive, that S is dependent if and only if some vector in S does depend on the others in the set. If S is dependent then there is a non trivial linear combination $0 = c_1v_1+....+c_rv_r$ in which at least on coefficient is non zero. Assume, after renumbering, the non zero coefficient is $c_1 \neq 0$. Then we can solve for $c_1v_1 = -c_2v_2 - c_3v_3 .....- c_rv_r$, and since c_1 is a non zero element of a field, we can divide by it, and get $v_1 = -(c_2/c_1)v_2 - (c_3/c_1)v_3 .....-(c_r/c_1)v_r$, which shows that v_1 depends on the other vectors $\{v_2,....,v_r\}$.
Conversely, if say v_1 depends on the other vectors $\{v_2,....,v_r\}$, by means of a linear combination $v_1 = a_2v_2 + a_3v_3 .....+a_rv_r$, then we can write the zero vector as
$0 = - v_1 + a_2v_2 + a_3v_3 .....+a_rv_r$. This is a non trivial linear combination since the first coefficient is $-1 \neq 0$, so the set $S = \{v_1,....,v_r\}$ is dependent.
(2) If S is a spanning set for V we will show S is dependent if and only if some proper subset of S does still span V. If S is dependent, then some vector in S, say v_1, depends on the others, e.g. (*) $v_1 = a_2v_2 + a_3v_3 .....+a_rv_r$ for some coefficients a_i. Then we claim the subset $\{v_2,....,v_r\}$ still spans V, since we can use the equation (*) to substitute for v_1 in any linear combination. I.e. if w is any vector in V, w depends on S, so there is a linear combination $w = c_1v_1+....+c_rv_r$. Then using (*) we get $w = c_1(a_2v_2 + a_3v_3 .....+a_rv_r) +c_2v_2+.....+c_rv_r$
$= c_1a_2v_2 + c_1a_3v_3 .....+c_1a_rv_r +c_2v_2+.....+c_rv_r$
$= (c_1a_2+c_2)v_2 + (c_1a_3+c_3)v_3+.....+(c_1a_r+c_r)v_r,$
which shows that w depends on the proper subset $\{v_2,....,v_r\}$.
Conversely, if some proper subset of S still spans, then some subset spans which is obtained by removing only one element, say v_1. But if $\{v_2,....,v_r\}$ still spans V, then in particular v_1 depends on this set by a linear combination $v_1 = a_2v_2 + a_3v_3+ ..... +a_rv_r$. Then $0 = -v_1 + a_2v_2 + a_3v_3 .....+a_rv_r$, so the set $\{v_1, v_2,....,v_r\}$ is dependent.
QED.

Definition: A subset $S = \{v_1,....,vr\}$ is called a <u>basis</u> of a vector space V over F, if and only if S is independent and spans V.

The main result that allows us to define dimension is the next one.

Theorem: Every vector space has a basis, and any two bases have the same number of elements.

Sketch of proof: It is not difficult to produce a basis of a space in case there is a finite spanning set. If our spanning set is independent, it is already a basis. If not, there is some vector in it that depends on the others, and it can then be removed and the remaining set still spans. If the remaining set is now independent, we have a basis. If not, we can remove another vector and still have a spanning set. If we ever get to an independent spanning set we have a basis. If not, we get down to a set of only one vector, which still spans but is not independent. Now if a vector v is non zero, then the set $\{v\}$ is independent. I.e. if $c \neq 0$, and $cv = 0$, then $v = 1v = (c^{-1}c)v) = c^{-1}(cv) = c^{-1} = 0$, so it is impossible to have a non trivial expression $cv = 0$, when $v \neq 0$. Thus when we get down to one vector, if it is non zero it is a basis, and if it is zero, since it spans V, then the space $V = \{0\}$. If the space is just $\{0\}$, we agree that the empty set is a basis. If $V \neq \{0\}$ is a non zero vector space with a finite spanning set, by this process we get a finite basis for V with ≥ 1 element in it.

The proof that any two bases have the same number of elements is more work, but it follows from the results of Math 3000. I.e. if $\{v_1,....,v_r\}$ and $\{w_1,....,w_s\}$ are two bases of V, then each vector v_j can be written in terms of the vectors $\{w_1,....,w_s\}$ using a sequence of scalars $(a_1,....,a_s)$. Viewing this sequence of scalars as a column in a matrix, the vectors $\{v_1,....,v_r\}$ are represented as the columns of an "s by r" matrix A. Then a linear combination $c_1v_1+.....+c_rv_r$, is equivalent to a linear combination of the columns of the matrix M, i.e. to a matrix product of M with the column vector $[c_1\ c_2\ .....c_r]$. Recall that a homogeneous system of linear equations has a non trivial solution if there are more unknowns than equations, i.e. if in our matrix M, we have $r > s$, then there is a non zero column vector $c = [c_1\ c_2\ .....c_r]$ such that $Ac = 0$. Thus if $r > s$ then the set $\{v_1,....,v_r\}$ would be dependent, a contradiction since it is a basis. Thus if $\{v_1,....,v_r\}$ and $\{w_1,....,w_s\}$ are both bases, then $r \leq s$. Doing the argument in the other order, we get $s \leq r$, so $r = s$. **End of proof sketch.**

More details.
If we form row vectors $V = (v_1,....,v_r)$ and $W = (w_1,....,w_s)$ made up of the vectors in our two sets, and if the set of w's spans, then there is an s by r matrix of numbers M such that $WM = V$. We claim if $r > s$, then the set of v's is dependent, i.e. there is a column vector X of numbers, not all zero, such that $VX = 0$ (= the zero vector). It would suffice to have such a column vector X (i.e. a column of numbers in our field, not all zero) with $MX = 0$ (= the zero column vector), since then we have $VX = (WM)X = W(MX) = W0 = 0$. But if $r > s$, then the s by r matrix M has more columns than rows, hence by Gaussian elimination, there is a column vector X of numbers, not all zero, such that $MX = 0$ (= the zero column vector). QED.

Definition: The dimension of a vector space V over F, is the number of vectors in a basis for V over F.

Thus if V is a vector space and $S = \{v_1,....,v_r\}$ is an independent spanning set for V, then the dimension of V is r.

Now let's get down to business and compute some dimensions of the vector spaces of interest to us, namely root fields of irreducible polynomials.

Preview of Root fields:

If f,g are relatively prime in Q[X], then they cannot have a common root in C = complex numbers.

If f,g are relatively prime in Q[X], then they are still relatively prime over any larger field (i.e. one containing Q).

If r is a complex root of an irreducible polynomial f in Q[X], then r cannot be a root of any polynomial in Q[X] of degree less than deg(f).

If r is a complex number which is a root of some polynomial in Q[X], then there is a unique monic irreducible polynomial f in Q[X] such that r is a root of f. Then Q(r) = the smallest subfield of C containing (both Q and) r, has finite dimension over Q equal to deg(f), and the field Q(r) is isomorphic to the modular ring Q[X]/(f). In fact $1, r, r^2, \ldots, r^{n-1}$ span (over Q) the ring Q[r] of "polynomials in r" where n = deg(f) and they are independent over Q. Moreover this ring Q[r] is a field, hence it equals the field Q(r). (Note that every subfield of C contains Q.)

If r is an element of C satisfying a monic irreducible polynomial f in Q[X], and if F is any subfield of C containing (both Q and) r, and if the dimension of F over Q is finite and equal to n, then deg(f) divides n.

4000/6000 Day 25 Root fields for polynomials

Consider the fields Q = the rational numbers, and C = the complex numbers. In between these two fields are infinitely many other fields. If r is any complex number, there is a smallest subfield Q(r) of C containing (Q and) r. Notice that it unnecessary to say a subfield of C contains Q, since every field must contain 1 and 0, and all sums 1+1+1.....+1 = n, and all negatives -n of elements it contains, and all quotients of elements by non zero elements, so every subfield of C must contain all numbers of form, n/m, where n, m are integers and $m \neq 0$, i.e. every subfield of C must contain all rational numbers.

There are two kinds of elements of C, the algebraic elements and the transcendental elements. An algebraic element of C is one that satisfies some non constant polynomial over Q. Thus all elements of Q are algebraic, since if r is rational it satisfies X-r = 0, and sqrt(2) is algebraic since it satisfies X^2-2. It is hard to identify transcendental numbers, those which do not satisfy any non constant polynomial equation over Q, but it has been proved that π and e are transcendental, as well as numbers like

.0100100000010000000000000000000000001....., where each 1 occurs after n! zeroes, for n = 1,2,3,.....

This proof is in our book, but is less interesting than the harder proofs for π, and e, for which one must look in a number theory book, or an honors calculus book like Spivak, used in 2400H here.

We will focus on algebraic numbers. If we forget some of the multiplicative structure, then every field containing Q can be viewed as a vector space over Q. The main result is the following:

Theorem: If r is a complex number, then r is algebraic over Q if and only if the field Q(r) is finite dimensional as a vector space over Q, and hence r is transcendental if and only if this field is infinite dimensional over Q.

The concept of vector dimension over Q is defined in terms of the concept of linear combinations with coefficients in Q.
The main example of vector space we want is a pair of fields E, F containing Q. I.e. if E contains F, which contains Q, then E is a vector space over F. The vector addition in E is the usual addition in E, and the multiplication of elements of E by elements of F is the multiplication in E, but you just forget about the multipliaction of elements of E where neither is in F. So E, considered as a vector space over F, has a little less structure than it has as a field.

Now consider any complex number r and a subfield F of C. We want to know what the smallest subring, and the smallest subfield of C containing r and F, look like.

Definition: The smallest subring of C containing r and the subfield F of C, is the intersection of all subrings of C containing F and r. It is the unique subring of C which contains r and which is contained in every other subring of C containing r and F. This ring is also called the subring of C generated by r over F.

Lemma: The smallest subring of C containing r and the subfield F, denoted F[r], consists of the

set of all numbers of form $a_0 + a_1r + a_2r^2 + ...+a_nr^n$, where the coefficients a_i are in F, and $n \geq 0$. I.e. it consists of all "polynomials" in r with coefficients in F.
proof: Since any ring containing F and r, contains all products and sums of things it contains, any ring containing r and F must contain all these elements. On the other hand, since any sum and any product of such polynomial expressions is again such a polynomial expression, this set is closed under sums and products. Moreover this set contains additive inverses of all its elements, hence is a ring. **QED.**

Definition: An element r of C is called "algebraic over (a subfield) F", if and only if it satisfies some non constant polynomial $a_0 + a_1X + a_2X^2 + ...+a_nX^n$, over F. I.e. if and only if there exist numbers $a_0, a_1,....,a_n$ in F, with $n \geq 1$, and $a_n \neq 0$, such that $a_0 + a_1r + a_2r^2 + ...+a_nr^n = 0$. A complex number which is algebraic over Q is simply called an "algebraic" number.

Lemma: Two polynomials f,g in F[X] which are relatively prime in F[X], cannot have a common root in any field containing F.
Proof: If f,g are relatively prime in F[X], there exist polynomials h,k over F such that $fh + gk = 1$. If r were a common root of f and g in any field containing F, then substituting X = r, would give 0 on the left but 1 on the right. This contradiction proves the result. **QED.**

The following result assigns to each algebraic number over F, a special irreducible polynomial with coefficients in F.

Proposition: If r in C is algebraic over a subfield F, then there is a unique monic irreducible polynomial over F satisfied by r.
Proof: First we show r satisfies some irreducible monic polynomial. Since r is algebraic, r satisfies some non constant polynomial over F, say g(r) = 0. Since g is non constant, and F[X] has unique factorization, we can factor g into irreducible factors, say $g = f_1f_2.....f_m$ where all f_i are irreducible over F. Then $0 = g(r) = f_1(r)f_2(r).....f_m(r)$, and since C is a field, one of the factors $f_i(r)$ must equal zero. Thus r is a root of the irreducible polynomial $f_i(X)$. Dividing by the leading coefficient c of f_i does not change this fact, so then r satisfies the monic polynomial f_i/c.

Now we show the monic irreducible polynomial satisfied by r is unique. Suppose f,g are two different monic irreducible polynomials. then neither can divide the other since if f = gu, then since f and g are irreducible u is a unit, hence u is a non zero element of F. But since both f,g are monic, the leading coefficient of f, which is 1, equals u times the leading coefficent of h, so u = 1, and hence we would have f = g. Thus if f and g are different monic irreducible polynomials over a subfield F of C, then they are relatively prime. Then they cannot have a common root. Thus an algebraic element r over F, is a root of exactly one monic irreducible polynomial over F. **QED.**

Definition: If r is any element of C, the smallest subfield of C containing r and a subfield F, is defined to be the intersection of all subfields of C containing r and F. it is the unique subfield of C containing r and F and contained in all other subfields which contain r and F. This field is denoted F(r) and is also called the subfield of C generated by r over F.

Theorem: If r in C is algebraic over a subfield F of C, then the ring F[r] is already a field, and hence F[r] = F(r) = the smallest subfield of C containing r and F.

proof: We only need show that every number of form $a_0 + a_1r + a_2r^2 + ...+a_nr^n$ has a multiplicative inverse of the same form, (but possibly of different "degree"). We know this already. I.e. suppose r satisfies the irreducible polynomial f over F. If g is any non zero polynomial over F, and g(r) the corresponding element of F[r], if g(r) is not zero, then f does not divide g. Then f and g are relatively prime, so there exist polynomials h,k over F such that fh+gk = 1. Setting X = r gives f(r) = 0, so g(r)k(r) = 1, and hence k(r) is the inverse of g(r). **QED.**

Lemma: If r in C is algebraic over a subfield F of C, and its monic irreducible polynomial f over F has degree $n \geq 1$, then the field F[r] is a vector space over F, and the elements $\{1, r, r^2,...., r^{n-1}\}$ form a basis over F. In particular, F[r] has dimension n over F.

Proof: To show it spans, let g be any polynomial over F and g(r) the corresponding element of F[r]. Then by the division algorithm,. we can find polyonmials h,k such that g = hf + k, and deg(k) < deg(f) = n. Then
g(r) = h(r)f(r) + k(r),a nd since f(r) = 0, g(r) = k(r). Since k has degree < n, k(r) has form $a_0+a_1r+a_2r^2+....+a_{n-1}r^{n-1}$, hence g(r) is a linear combination of 1, r, $r^2,....,r^{n-1}$ as claimed.

To show independence, note that if $\{1, r, r^2,...., r^{n-1}\}$ were dependent, there would be coefficients $c_0,....,c_{n-1}$, not all zero, such that

$c_0 + c_1r+ c_2 r^2+.....+c_{n-1} r^{n-1} = 0$. Then r satisfies the non zero polynomial

$g(X) = c_0 + c_1X+ c_2 X^2+.....+c_{n-1} X^{n-1}$. But since g is non zero and of lower degree than f, and f is irreducible, then f and g are relatively prime. But we have proved then that f,g can not have any common roots, so since r satisifes f, this is a contradiction.

QED.

Corollary: The field Q(sqrt(2)) has dimension 2 over Q, and Q(cuberoot(2)) has dimension 3.

Proof: The element sqrt(2) satisfies $X^2 - 2 = 0$, and cuberoot(2) satisfies $X^3 - 2 = 0$, and these are both irreducible. **QED.**

4000/6000 Day 26. Dimension of algebraic field extensions (root fields)

The fundamental result, proved in the day 25 notes, is that the dimension of the field F[r] obtained by adjoining an algebraic element r to a field F, equals the dergee of any irreducible monic polynomial satisfied by r over F. We will prove this again today, and derive as a consequence that if $r = \cos(20^0) = \cos(\pi/9)$, then Q[r] has dimension 3 over Q.

The connection between polynomials and linear combinations is based on the following result.

I). The polynomial ring F[X] as a vector space over F.

If F is any field, then the polynomial $a_0 + a_1X + a_2X^2 + ...+a_nX^n$ is precisely a linear combination of the elements $(1, X, X^2,.....,X^n)$ with coefficients $a_0, a_1, a_2, ..., a_n$. Since the only way for f to be the zero, polynomial is for all coefficients a_i to be zero, this says that the elements
$(1, X, X^2,.....,X^n)$ of F[X] are independent, for every $n \geq 0$. Since for every $n \geq 1$, F[X] contains an independent set of n elements, by problem 24, section 5.1, there can be no finite spanning set in F[X]. Thus F[X] is infinite dimensional over F.

Next we examine the connection between a root field and polynomials.

II). Let F be any subfield of the complex numbers C and let r be any complex number. Then the numbers $(1,r,r^2,r^3,....,r^n)$ are linearly dependent over F, if and only if r satisfies some non zero polynomial over F of degree $\leq n$.

Proof: If r satisfies the non zero polynomial $a_0 + a_1X + a_2X^2 + ...+a_nX^n$, it means that $a_0 + a_1r + a_2r^2 + ...+a_nr^n = 0$, but that some of the coefficients a_i are not zero. This is exactly what it means to say that the set $(1,r,r^2,r^3,....,r^n)$ is dependent over F. I.e. conversely, if $a_0 + a_1r + a_2r^2 + ...+a_nr^n = 0$, is a non trivial lienar combination of the elements $(1,r,r^2,r^3,....,r^n)$ which equals zero, then the polynomial $a_0 + a_1X + a_2X^2 + ...+a_nX^n$, is non zero of degree $\leq n$, and is satisfied by r. **QED.**

III). Recall if r is a complex number, abd F a subfield of C, then F[r] = the smallest subring of C containing both F and r = all complex numbers of form $a_0 + a_1r + a_2r^2 + ...+a_nr^n$ for all $n \geq 0$, and all coefficients a_i in $F = \text{span}(1,r,r^2,r^3,.....)$ (an infinite sequence of powers of r).

Lemma: If r is algebraic over F, and is a root of a non zero polynomial $a_0 + a_1X + a_2X^2 + ...+a_nX^n$, of degree $n \geq 0$, then F[r] is spanned as a vector space over F by the set of n elements $(1,r,r^2,r^3,....,r^{n-1})$. In particular F[r] has finite dimension $\leq n$ over F.

Proof: This follows from the division theorem. I.e. we must show any element g(r) of F[r], where g is any polynomial over F, depends on the set $(1,r,r^2,r^3,....,r^{n-1})$. By the division theorem, we get g = fh + k, where h,k, are in F[X] and either k = 0 or deg(k) < deg(f) = n. Then substituting X = r, we get g(r) = f(r)h(r) + k(r) = k(r), since f(r) = 0. And since either k = 0 or

$\deg(k) \leq n-1$, $k(r) = c_0 + c_1 r + c_2 r^2 + \ldots + c_{n-1} r^{n-1}$, for some elements c_i in F. Thus g(r) depends on $(1,r,r^2,r^3,\ldots,r^{n-1})$ as claimed. **QED.**

IV). If r is a complex number, F a subfield of C, and if r satisfies an irreducible polynomial f over F of degree $n \geq 0$, then in fact $n \geq 1$ and the set $(1,r,r^2,r^3,\ldots,r^{n-1})$ is a basis of F[r] over F, in particular F[r] has dimension n over F.
Proof: A non zero polynomial of degree 0 is a non zero constant, hence has no roots, so if r satisfies f, we must have $\deg(f) = n \geq 1$. By III, the set $(1,r,r^2,r^3,\ldots,r^{n-1})$ spans F[r], so we must show it is independent. If it were dependent, then by II, r would satisfy a non zero polynomial g over F of degree $\leq n-1$. But since f is irreducible of degree n, then f and g would be relatively prime, and we already know from last time that then they cannot have a common root r. Hence then $(1,r,r^2,r^3,\ldots,r^{n-1})$ is a basis of F[r], which thus has dimension n over F. **QED.**

Notation: For this reason, in field theory it is usual to call the dimension of a field extension, the "degree" of the extension, and to write it differently. I.e. if E is a field containing another field F, then the dimension of E over F as a vector space, is called the "degree of E over F", and is written as [E : F] = degree (i.e. dimension) of E over F.

V). If $r = \cos(20^0) = \cos(\pi/9)$, then Q[r] has dimension 3 over Q.
Proof: It suffices to find an irreducible polynomial of degree 3 over Q which is satisfied by r. We use the trick of complex exponentials, which gave us the double angle formula, to give us a triple angle formula. I.e. we know that $\cos(\pi/3) = \cos(60^0) = 1/2$, by looking at an equilateral triangle, cut in half. Thus we want to express the number $\cos(\pi/3) = 1/2$, as a cubic polynomial in the number $r = \cos(\pi/9)$. Recall that $e^{i\pi/3} = \cos(\pi/3) + i\sin(\pi/3) = 1/2 + i\,\mathrm{sqrt}(3)/2$. And also that $e^{(ab)} = (e^a)^b$, so since $\pi/3 = 3\pi/9$ we have $e^{i\pi/3} = e^{i(3\pi/9)} = [e^{i\pi/9}]^3$
$= [\cos(\pi/9) + i\sin(\pi/9)]^3 =$
$\cos^3(\pi/9) + 3\cos^2(\pi/9) i \sin(\pi/9) + 3\cos(\pi/9) i^2 \sin^2(\pi/9) + i^3 \sin^3(\pi/9)$

$= \cos^3(\pi/9) + 3i\cos^2(\pi/9)\sin(\pi/9) - 3\cos(\pi/9)\sin^2(\pi/9) - i\sin^3(\pi/9)$, and if we write $r = \cos(\pi/9)$, and $s = \sin(\pi/9)$, this becomes

$= r^3 + 3ir^2 s - 3rs^2 - is^3 = (r^3 - 3rs^2) + i(3r^2 s - s^3)$

$= 1/2 + i\,\mathrm{sqrt}(3)/2.$

Thus we have $(r^3 - 3rs^2) = 1/2$.

Since also $r = \cos(\pi/9)$ and $s = \sin(\pi/9)$, we have $r^2 + s^2 = 1$, so $s^2 = 1 - r^2$, so substituting gives $(r^3 - 3r(1-r^2) = (4r^3 - 3r) = 1/2$.

Thus r satisfies the equation $4X^2 - 3X - (1/2) = 0$, or $8X^3 - 6X - 1 = 0$.

We want to show that $8X^3 - 6X - 1$ is irreducible over Q, and it suffices to show it has no factors

of degree one, or equivalently no rational roots. This can be checked in a few minutes, and involves trying the 8 possibilities X = 1, -1, 1/2, -1/2, 1/4, -1/4, 1/8, and -1/8. It would also suffice to find a prime integer p such that p does not divide 8, and the reduced polynomial over Z_p has no roots. We cannot reduce mod 2 since that lowers the degree of then polynomial. Also the test fails for p = 3, since the reduced polynomial mod 3, is $-X^3-1$, which does have X = -1 as a root. Mod p = 5 however, the polynomial becomes $h(X) = 3X^3 - X - 1$, and we only have to try 0, 1, -1, 2, -2, and none of these is a root, since they give h(0) = -1, h(1) = 1, h(-1) = -3, h(2) = 1, and h(-2) = -3. [Please check me on this. If any one of these numbers should really be zero, the whole calculation goes down the drain and the result could be false.] **QED.**

VI). We also want to know how to compute the dimension of larger fields, obtained by adjoining several new elements, such as the dimension of F[r,s] over F, where r,s, are complex numbers, and F a subfield of C. Unfortunately, simply knowing the dimensions of the two fields F[r] and F[s] over F, does not always determine the dimension of F[r,s] over F. There is one acse where it does, in an example in the book. We will prove that if the two dimensions are relatively prime, i.e. if say [F[r]: F] = n, and [F[s] : F] = m, where n and m are relatively prime, then [F[r,s] : F] = mn. The general result is that if F, E, K are subfields of C and E contains F, and K contains E, then the degree of E over F, times the degree of K over E, equals the degree of K over F.

Here is an example: consider $X^4 - 2$ over Q, an irreducible quartic polynomial. It has 4 complex roots, $\sqrt[4]{2}$ (= the real positive 4th root of 2), $-\sqrt[4]{2}$, $i\sqrt[4]{2}$, and $-i\sqrt[4]{2}$. By what we have proved today, adjoining any one of these roots to Q gives a field of dimension 4 over Q. But depending on which root you adjoin, you get two different fields. I.e. adjoining $\sqrt[4]{2}$ gives the same field as adjoining $-\sqrt[4]{2}$, and this field is contained in the real numbers. Thus $Q[\sqrt[4]{2}]$ is a subspace of R which is 4 dimensional over Q. This field contains only real numbers hence does not contain the roots $i\sqrt[4]{2}$, and $-i\sqrt[4]{2}$. On the other hand, the field $Q[i\sqrt[4]{2}]$ is also 4 dimensional over Q, but is not contained in R. It also contains $-i\sqrt[4]{2}$, but not either root $\sqrt[4]{2}$ or $-\sqrt[4]{2}$ (which is not as obvious, except by counting dimensions).

Thus if we want to adjoin all the complex roots of $X^4 - 2$, we can do it in two stages, first adjoin $\sqrt[4]{2}$, and then adjoin also $i\sqrt[4]{2}$. Now since $X^4 - 2$ is irreducible over Q, the field $Q[\sqrt[4]{2}]$ has dimension 4 over Q, as we said. But then over the larger field $Q[\sqrt[4]{2}]$, $X^4 - 2$ is no longer irreducible by the root factor theorem, since it has a root in the field. For instance, we can factor it easily recalling that 2 is now a square in this field, since it is the square of $(\sqrt[4]{2})^2 = \sqrt[4]{4}$.

Then we get $X^4 - 2 = (X^2)^2 - (\sqrt[4]{4})^2 = (X^2 - \sqrt[4]{4})(X^2 + \sqrt[4]{4})$, and now

since also $\sqrt[4]{4}$ is a square, namely $\sqrt[4]{4} = (\sqrt[4]{2})^2$, we also get

$X^4 - 2 = (X - \sqrt[4]{2})(X + \sqrt[4]{2})(X^2 + \sqrt[4]{4})$ in the field $Q[\sqrt[4]{2}]$.

Now in this field the last quadratic, $(X^2 + \sqrt[4]{4})$, is still irreducible, since it has negative discriminant, hence no real roots.

Thus over the field $Q[\sqrt[4]{2}]$, the complex 4th root $i\sqrt[4]{2}$ has degree two. We will show that the field $Q[\sqrt[4]{2}, i\sqrt[4]{2}] = Q[\sqrt[4]{2}, i]$, has degree 8 over Q.

Lemma: If F is a subfield of E, and E is a subfield of K, then $[K : F] = [K : E][E : F]$. In fact, if $(z_1,....,z_n)$ is a basis of E over F, and if $(w_1,....,w_m)$ is a basis of K over E, then the set of nm pairwise products $(\{z_i w_j\})$, for all $1 \leq i \leq n$, $1 \leq j \leq m$, is a basis of K over F.

Proof: This is just a trivial computation, using distributivity. Anyone can learn to do it. But since it is a little messy, I want to illustrate it first with some small numbers, i.e. say $n = [E : F] = 2$, and $m = [K : E] = 3$, and the bases are (z_1, z_2) of E over F, and (w_1, w_2, w_3) of K over E. Then look at the set

$\{z_1w_1, z_1w_2, z_1w_3, \quad z_2w_1, z_2w_2, z_2w_3\}$, we claim is a basis of K over F.

To show it spans K over F, let u be any element of K. Since (w_1, w_2, w_3) spans K over E, we can write $u = b_1w_1 + b_2w_2 + b_3w_3$ for some b_j in E.

Then since (z_1, z_2) spans E over F, each b_j can be written in terms of (z_1, z_2) using coefficients in F, say $b_1 = a_{11}z_1 + a_{21}z_2$,

$b_2 = a_{12}z_1 + a_{22}z_2$, and $b_3 = a_{13}z_1 + a_{23}z_2$.

Then $u = b_1w_1 + b_2w_2 + b_3w_3$

$= (a_{11}z_1 + a_{21}z_2)w_1 + (a_{12}z_1 + a_{22}z_2)w_2 + (a_{13}z_1 + a_{23}z_2)w_3$.

Now we can expand this as a linear combination of the elements

$\{z_1w_1, z_1w_2, z_1w_3, \quad z_2w_1, z_2w_2, z_2w_3\}$, with coeffificients a_{ij} in F.

I.e.

$u = (a_{11}z_1 + a_{21}z_2)w_1 + (a_{12}z_1 + a_{22}z_2)w_2 + (a_{13}z_1 + a_{23}z_2)w_3$

$= a_{11}(z_1w_1) + a_{21}(z_2w_1) + a_{12}(z_1w_2)$

$+ a_{22}(z_2w_2) + a_{13}(z_1w_3) + a_{23}(z_2w_3)$. Since we have expressed an arbitrary element u of K as

a linear combination of the elements
$\{z_1w_1, z_1w_2, z_1w_3, \quad z_2w_1, z_2w_2, z_2w_3\}$ with coefficients in F, these elements span K over F.

To see the set $\{z_1w_1, z_1w_2, z_1w_3, \quad z_2w_1, z_2w_2, z_2w_3\}$ is independent over F, assume we have a linear combination equalling zero.

$0 = a_{11}(z_1w_1) + a_{21}(z_2w_1) + a_{12}(z_1w_2)$

$+ a_{22}(z_2w_2) + a_{13}(z_1w_3) + a_{23}(z_2w_3).$

We must show the coefficients a_{ij} are all zero. We work backwards in comparison to what we just did before. I.e. expand as a linear combination involving of the w_j,

$0 = a_{11}(z_1w_1) + a_{21}(z_2w_1) + a_{12}(z_1w_2)$

$+ a_{22}(z_2w_2) + a_{13}(z_1w_3) + a_{23}(z_2w_3)$

$= (a_{11}z_1 + a_{21}z_2)\, w_1 + (a_{12}z_1 + a_{22}z_2)\, w_2 + (a_{13}\, z_1 + a_{23}\, z_2)\, w_3.$

Now observe all these new coefficients involve only a's and z's, so since the a's belong to F and the z's belong to E, all these coefficients belong to E. Since by hypothesis, the wj are independent over E, all these coefficients must be zero. I.e.l this gives us three linear combinations, all equal to zero.

$(a_{11}z_1 + a_{21}z_2) = 0,$

$(a_{12}z_1 + a_{22}z_2) = 0,$ and

$(a_{13}\, z_1 + a_{23}\, z_2) = 0.$

These equations are linear combinations of the z_i's, with coefficients in F. Since the z_i's are independent over F, each coefficient is zero, i.e. every $a_{ij} = 0$, as desired. Thus in fact the set $\{z_1w_1, z_1w_2, z_1w_3, \quad z_2w_1, z_2w_2, z_2w_3\}$, is a basis of K over F. **QED.**

Now we give the shorter, but possibly more opaque general proof.\

Let $(z_1, \ldots, z_n)$ be a basis of E over F, and $(w_1, \ldots, w_m)$ a basis of K over E. Then we claim the set of pairwise products $(\{z_iw_j\})$, for $1 \leq i \leq n$, $1 \leq j \leq m$, is a basis of K over F.

spanning: Given any element u of K, it can be expressed as $u = \sum_j b_j\, w_j$, where the b_j are in E, and the sum runs from j =1, to j = m.

Each b_j in turn can be expressed as (*) $b_j = \sum_i a_{ij} z_i$, where the a_{ij} are in F, and the sum runs from i = 1, to i = n.

Then substituting the equations (*) for the b_j into the first equation for u, gives us

$u = \sum_j b_j w_j = \sum_j (\sum_i a_{ij} z_i) w_j = \sum_{i,j} a_{ij} (z_i w_j)$, which expresses u as a linear combination of the desired products with coefficients in F.

independence:
To show independence we again work backwards, assuming that

$0 = \sum_{i,j} a_{ij} (z_i w_j) = \sum_j (\sum_i a_{ij} z_i) w_j$, which implies, by the independence of the w_j over E, that for every j, we have the linear combination

$\sum_i a_{ij} z_i = 0$. Then by the independence of the z_i over F, this implies for every j, that all the coefficients a_{ij} are zero. This does it. I.e. the set ($\{z_i w_j\}$), for $1 \leq i \leq n$, $1 \leq j \leq m$, is both linearly independent and spanning for K over F, hence it is a basis of K over F. **QED.**

Here is a matrix version of the proof of this theorem: Let Z be the row vector whose entries are the numbers $(z_1,....,z_n)$, and let W be the column vector whose entries are the numbers $(w_1,....,w_m)$. Then if A is an n by m matrix of entries from F, the product ZAW is exactly a sum of form $\sum_{i,j} a_{ij} (z_i w_j)$. Hence to show the products $(z_i w_j)$ span K over F, means, given an arbitrary element u in K, to find an n by m matrix A over F such that ZAW = u. Now since the w's span K over E, there is a row vector B of length m such that BW = u. Then since the z's span E over F, there is an n by m matrix A over F such that ZA = B. Then u = BW = (ZA)W, as desired. To show independence, means that if A is any n by m matrix over f such that ZAW = 0, then A = 0. But if ZAW = 0, then (ZA)W = 0, so by independence of the w's over E, since the entries of ZA belong to E, every entry of the row vector ZA is zero. But by independence of the z's over F, then every column of A consists of all zeroes. **QED.**

I think the double summation version is easier, but I had trouble making it look easy to my class.

The following corollaries are of interest to us.

Corollary: If r is a complex number which satisfies an irreducible polynomial over Q of degree n, and is F is any subfield of C containing r and finite dimensional over Q, then n must divide the dimension [F : Q].

Proof: We just showed that $[F : Q] = [F : Q[r]]\,[Q[r] : Q]$, and we know that $[Q[r] : Q] = n$. **QED.**

Corollary: If r,s are two elements of C, if $[Q[r] : Q] = n$, and if $[Q[s] : Q] = m$, and if n,m are relatively prime, then $[Q[r,s] : Q] = nm$.
Proof: Let Q[r] be the smallest subfield of C containing r, and let (Q[r])[s] be the smallest subfield containing the field Q[r] and the element s. Then this is the same as Q[r,s] = the smallest subfield of C containing both r and s.

Hence the dimension of Q[r,s] over Q[r] equals the degree of the lowest degree polynomial satisifed by s with coefficients in Q[r]. Since the irreducible polynomial of s over Q is one such polynomial, but not necessarily of lowest possible degree, since it only allows coefficients in Q, we learn that $[Q[r,s] : Q[r]\,] \leq [Q[s] : Q] = m$. Thus $[Q[r,s] : Q]$
$= [Q[r,s] : Q[r]\,]\,[Q[r] : Q\,] \leq [Q[s] : Q]\,[Q[r] : Q\,] = nm$.

On the other hand, both n and m must divide $[Q[r,s] : Q]$. Since n,m are relatively prime their product also divides $[Q[r,s] : Q]$, and then the inequality implies that $[Q[r,s] : Q] = nm$. **QED.**

4000/6000 Day 27 Calculations with field extensions

Explicit construction of fields.
We will construct all possible finite dimensional fields containing Q, i.e. all fields containing Q and finite dimensional as vector space over Q.
Let μ be a symbol, and choose a dimension for your field, say n. Then let the basis elements of your field be given by the symbols $1,\mu,\mu^2,....,\mu^{n-1}$. Then define the field to consist of all expressions of form
$a_0 + a_1\mu + a_2\mu^2+....+a_{n-1}\mu^{n-1}$, where the coefficients a_i are any rational numbers. Then add two of these as usual, i.e.
$(a_0 + a_1\mu + a_2\mu^2+....+a_{n-1}\mu^{n-1}) + (b_0 + b_1\mu + b_2\mu^2+....+b_{n-1}\mu^{n-1})$

$= (a_0+b_0) + (a_1+b_1)\mu + (a_2+b_2)\mu^2+....+(a_{n-1}+b_{n-1})\mu^{n-1}$.

If c is any rational number, multiply also in the usual way, i.e.
$c\,(a_0 + a_1\mu + a_2\mu^2+....+a_{n-1}\mu^{n-1}) = (ca_0 + ca_1\mu + ca_2\mu^2+....+ca_{n-1}\mu^{n-1})$.

So far all we have done is define the vector space structure, and we have given it the usual vector space sructure of the rational n dimensional space Q^n. I.e. the element $(a_0 + a_1\mu + a_2\mu^2+....+a_{n-1}\mu^{n-1})$ of our field, can be thought of as corresponding to the vector of coefficients $(a_0, a_1, ...,a_{n-1})$ which is just a point of Q^n.

But now we want to define a multiplication on our vector space so it really becomes a field. I.e. we have to say how to multiply two elements

$(a_0 + a_1\mu + a_2\mu^2+....+a_{n-1}\mu^{n-1})$ and $(b_0 + b_1\mu + b_2\mu^2+....+b_{n-1}\mu^{n-1})$
of our field together and get for an answer another element like that

i.e. like $(c_0 + c_1\mu + c_2\mu^2+....+c_{n-1}\mu^{n-1})$, where the c_i are rational numbers. And all the rules for rings and fields have to be true.

Well we start off by just multiplying as if they were polynomials, but we get terms of degree higher than n-1 in our symbol μ. E.g. if n = 3, we have

$(a_0 + a_1\mu + a_2\mu^2)(b_0 + b_1\mu + b_2\mu^2) =$

$a_0b_0+(a_0b_1+a_1b_0)\mu +(a_0b_2+a_1b_1+a_2b_0)\mu^2 + (a_1b_2+a_2b_1)\mu^3 +(a_2b_2)\mu^4$.

Now we have to decide what μ^3 and μ^4 should be equal to. I.e. they have to be set equal to some expressions of degree less than or equal to 2, in μ.

This can be done in many ways, by imitating the construction of a modular polynomial ring,

modded out by an irreducible polynomial.

I.e. all we need is any irreducible polynomial f(X) of degree n over Q, in this case of degree 3, such as $f(X) = X^3 - p$, where p is a prime integer, or $f(X) = X^3-X-1$, or $f(X) = X^3 + 29X^2 - 58X + 29$ (by Eisenstein), or ANY other irreducible polynomial over Q, of degree n.

Then we use this polynomial to tell us how to redefine μ^n. I.e. assume μ is a root of our polynomial, so we set our polynomial, with μ substituted for X, equal to zero.

In case n = 3, and our polynomial is $f(X) = X^3-X-1$, we set $\mu^3 - \mu - 1 = 0$, which just means that $\mu^3 = \mu+1$. Then we get $\mu^4 = \mu(\mu^3) = \mu^2+\mu$. Thus in the expression above for a product, we have

$(a_0 + a_1\mu + a_2\mu^2)(b_0 + b_1\mu + b_2\mu^2) =$

$a_0b_0+(a_0b_1+a_1b_0)\mu +(a_0b_2+a_1b_1+a_2b_0)\mu^2 + (a_1b_2+a_2b_1)\mu^3 +(a_2b_2)\mu^4$

$= a_0b_0+(a_0b_1+a_1b_0)\mu +(a_0b_2+a_1b_1+a_2b_0)\mu^2 + (a_1b_2+a_2b_1)(\mu+1) +(a_2b_2)(\mu^2+\mu).$

Now just multiply through and simplify to get the formula for the product.
Another way to say it is to take any product $A(\mu)B(\mu) = g(\mu)$, where A,B are polynomials over Q, and reduce it mod f, i.e. divide to get $g(X) = f(X)q(X) + r(X)$, where $\deg(r) < \deg(f)$ (or $r = 0$), and then replace $g(\mu)$ by $r(\mu)$, which will have degree < n as desired.

This is a field since it has the same structure as the modular ring Q[X]/(f), which we know is a field when f is irreducible. There is also a theorem that all finite dimensional fields, even if we construct them by adding several elements to Q, could have been constructed instead by adding one suitably chosen element. Hence all finite dimensional fields have the form of Q[X]/(f). (For example, the field Q[sqrt(2), i], obtained by adjoining to Q both sqrt(2) and i, could have been obtained by adjoining the one element sqrt(2)+i instead.)

Examples of field extensions

I. Let Q be the rational field and then $X^2 - 2$ is irreducible over Q, so if μ = sqrt(2) is the positive square root of 2, the smallest subfield $Q[\mu]$ of the complex field C, containing μ has vector dimesnion 2 over Q. In particular there is a Q - basis for $Q[\mu]$ consisting of the 2 numbers 1, μ. I.e. the elements of the field $Q[\mu]$ are exactly the numbers of form $a+b\mu$, where a,b are rational. (This says that $1,\mu$ span $Q[\mu]$ over Q.) Moreover $a+b\mu = c+d\mu$, iff a = c, and b = d. (Iff is not a misspelling - remember Sam's suggestion?) (This says that $1,\mu$ are independent over Q.)

The addition of elements is easy, since $(a+b\mu) + (c+d\mu) = (a+c) + (b+d)\mu$.

The multiplication is not hard either but you must remember that $\mu^2 = 2$. Thus $(a+b\mu)(c+d\mu) = ac + (ad+bc)\mu +bd\mu^2 = ac + (ad+bc)\mu +2bd$

$= (ac + 2bd) + (ad+bc)\mu$.

Division is a little harder. Recall it is done by rationalizing a denominator. I.e. since we know $(a+b\mu)(a-b\mu) = a^2 - b^2\mu^2 = a^2 - 2b^2$, and this is a rational number which is only zero when a= b =0, we get if $(a+b\mu) \neq 0$, then $1/(a+b\mu) = (a-b\mu)/[(a+b\mu)(a-b\mu)] = (a-b\mu)/(a^2 - 2b^2) = a/(a^2 - 2b^2) - \mu b/(a^2 - 2b^2) =$

$r + s\mu$, where $r = a/(a^2 - 2b^2)$ and $s = -b/(a^2 - 2b^2)$ are rational numbers.

II. Our second example was the field Q[i] obtained by adjoining a root i of the irreducible polynomial $X^2 + 1$ to Q. Here i denotes the root located in the upper half of the complex plane. Then the field Q[i] is again a 2 dimesnional vector space over Q, with basis, 1,i over Q, and hence elements uniquely expressible as a + bi where a,b, are rational. To multiply we need only multiply as usual and remember that $i^2 = -1$. The multiplicative inverse of a+bi is $a/(a^2+b^2) - i\, b/(a^2+b^2)$.

III. If we adjoin the real root $\mu = 2^{(1/5)}$, of the irreducible polynomial $X^5 - 2$, to Q we get a field Q[μ] of dimension 5 over Q, and basis $1,\mu,\mu^2, \mu^3, \mu^4$. I.e. every element can be written uniquely as $a+b\mu+c\mu^2+d\mu^3+e\mu^4 = a+b\sqrt[5]{2} +c\sqrt[5]{4} +d\sqrt[5]{8} +e\sqrt[5]{16}$. For inverse, if we ask for the inverse of f(r), where g is a polynomial of degree ≤ 4, then we can use the Euclidean algorithm to solve for polynomials h,k, such that $(X^5 - 2)h(X) + g(X)k(X) = 1$. Then setting $X = \mu = \sqrt[5]{2}$, gives the inverse of $g(\sqrt[5]{2})$ as $k(\sqrt[5]{2})$.

Applying this to $g(\sqrt[5]{2}) = 1+\sqrt[5]{2}$, which corresponds to g(X) = X+1, I got

$(X+1)(X^4 - X^3+X^2 - X +1) - 3 = X^5 - 2$, or

$(X+1)(X^4 - X^3+X^2 - X +1) - (X^5 - 2) = 3$, so it seems that

$(1+\sqrt[5]{2})^{-1} = [1-\sqrt[5]{2} + \sqrt[5]{4} -\sqrt[5]{8} +\sqrt[5]{16}]/3$. (Is it right? I did not check it! Check it and get to correct me!)

IV. What if we adjoin two elements, μ = sqrt(2) and i = sqrt(-1) to Q? We claim we get a vector space of dimension 4 over Q. I.e. adjoining μ to Q as we saw, gives a 2 dimensional extension Q[μ]. Then we must consider whether i is already in the field Q[μ]. But since μ is a real number, all elements $a+b\mu$, with a,b, rational, of the field Q[μ] are also real. So i does not belong to this field. (If you do not believe that kind of reasoning, and want more proof, ask yourself whether $(a+b\mu)^2 = a^2+2b^2 +2ab\mu$, can be -1. It cannot since all terms are non negative real

numbers.)

Since X^2+1 is quadratic and has no root in $Q[\mu]$, it is irreducible over $Q[\mu]$. Thus adjoining i to $Q[\mu]$ to get $Q[\mu,i]$, gives a vector which is 2 dimensional over $Q[\mu]$, and 1, i are a basis over $Q[\mu]$. Thus every element of $Q[\mu,i]$ can be written uniquely as $(a+b\mu) + (c+d\mu)i = a + b\mu + ci + d\mu i$, where a,b,c,d are rational Hence as a vector space over Q, it is spanned by $1,\mu,i, \mu i$. Thus the dimension over Q is at most 4, and it is exactly 4 if these elements are independent over Q,. But if there is some linear relation $a+b\mu + ci + d\mu i = 0$, then
$a+b\mu = - ci - d\mu i = (-c-d\mu)i$, and thus the purely real number on the left equals the "pure imaginary" number on the right. This means both are zero, so $a+b\mu = 0$, hence $a = b = 0$, since 1 and μ are independent over Q. Thus also $(c+d\mu)i = 0$, and since C is a field and $i \neq 0$, then $(c+d\mu) = 0$, so again by independence of $1,\mu$, we get $c = d = 0$.

In fact we know that degree of field extensions is always multiplicative, so that the dimension of $Q[\mu,i]$ over Q must be 4, and hence any spanning set of 4 elements must be a basis. (It at least contains a basis, but since all bases have exactly 4 elements it must itself be a basis.)

What about inverses in $Q[\mu,i]$? We know the inverse of elements of form $a+b\mu$, and also of elements of form $c+di$, but what about more complicated elements like $a+d\mu i$, or $a+b\mu+ci+d\mu i$?

Well we can play around with $a+d\mu i$, say multiply by its "conjugate" and get $(a+d\mu i)(a-d\mu i) = a^2 - d^2\mu^2 i^2 = a^2 + d^2\mu^2 = a^2 + 2d^2$, which is rational, and only zero when $a = d = 0$. Thus the inverse of $a+d\mu i$ is
$(a-d\mu i)/(a^2 + 2d^2)$, so this was as easy as the quadratic case.

As for the inverse of $a+b\mu+ci+d\mu i$, I do not know the answer to this offhand. Oh I guess I do. In fact I have two ways to do this one.

First way: do the two previous rationalization processes separately.

I.e. given $a+b\mu+ci+d\mu i$, rewrite it as $(a+b\mu)+i(c+d\mu)$, and then the inverse is $[(a+b\mu)-i(c+d\mu)]/[(a+b\mu)^2 + (c+d\mu)^2]$. Now all we have to do is get the μ's out of the bottom. But we can rewrite the bottom by multiplying it out, as something like $e+f\mu = a^2+2b^2 + 2ab\mu + c^2+2d^2 + 2cd\mu$

$= (a^2+2b^2 + c^2+2d^2) + (2cd + 2ab)\mu = e+f\mu$. So the inverse of this is

$(e-f\mu)/(e^2-2f^2) =$
$[(a^2+2b^2 + c^2+2d^2) - (2cd + 2ab)\mu]/[(a^2+2b^2+c^2+2d^2)^2 -2(2cd +2ab)^2]$.

Thus the inverse of $a+b\mu+ci+d\mu i$, should be

$[(a+b\mu)-i(c+d\mu)][(a^2+2b^2+c^2+2d^2)-(2cd + 2ab)\mu]$ divided by

$[(a^2+2b^2+c^2+2d^2)^2 -2(2cd +2ab)^2]$. At least this has no μ's or i's in the bottom.

Second way: We know a general method for finding inverse of fields of form Q[r] where r is algebraic over Q, so if we could find just one element r of Q[μ,i] such that Q[m,i] = Q[r], and the irreducible polynomial satisfied by r over Q, we could use our other mehthod, with the Euclidean algorithm to find inverses.

So let's try r = μ+i = sqrt(2) + i. At least this is contained in Q[μ,i]. So we ask whether both μ and i are contained in Q[r] = Q[μ+i] = the smallest subfield of C containing μ+i.

Well this field is closed under multiplication and division, addition and subtraction, so if it contains μ+i it must also contain $(\mu+i)^2 = 2 + 2\mu i - 1 = 1+2\mu i$. Since any field contains 1 and 2, it also contains μi. Then it also contains $\mu i(\mu+i) = 2i-\mu$. Then adding μ+i gives also 3i, hence i, hence also μ. So the field Q[μ+i] both contains and is contained in the field Q[μ,i], so they are equal. Now since the dimension of Q[μ,i] = Q[μ+i] over Q is known to be 4, there must be some irreducible polynomila of degree 4 with rational coefficients, satisfied by μ+i.

Let's try to find it. Well $(\mu+i)^2 = 2 - 1 + 2\mu i = 1+2\mu i$, so $(\mu+i)^2 - 1 = 2\mu i$, so $[(\mu+i)^2 - 1]/2 = \mu i$. Now $(\mu i)^2 = -2$, so $\{[(\mu+i)^2 - 1]/2\}^2 = -2$. Thus μ+i satisfies the polynomial $\{[X^2 - 1]/2\}^2 = -2$, i.e. $[X^2 - 1]^2/4 + 2 = 0$.

Thus, μ+i satisfies $[X^2 - 1]^2 + 8 = 0$. Since this has degree 4, and the field extension Q[μ+i] is known to have degree 4, then μ+i cannot satisfy any polynomial of degree lower than 4. Also any polynomial of degree 4 it satisfies must be irreducible, so this must be it. In particular $[X^2 - 1]^2 + 8 = X^4 - 2X^2 + 9$, must be irreducible over Q, even though i don ot know how to check that easily.

Well, I finally checked it mod 5 as follows. It reduces mod 5, to $X^4 + 3X - 1$, and this has no roots mod 5. Thus it cannot have linear factors mod 5, hence has no linear factors originally. To show it ahs no quadratic factors originally, I showed it has no quadratic factors mod 3 as follows. Mod 3 it reduces to $X^4 + X^2 = X^2(X^2+1)$ where X^2+1 is irreducible mod 3. Thus if it had two irreducible quadratic factors originally, one of them would reduce to X^2, and one of them to X^2+1, mod 3, so they would look like $(X^2+3aX+1)(X^2+3bX+9)$ originally, since the constant term was 9, and one factor reduced mod 3 to 1. But multiplying out $(X^2+3aX+1)(X^2+3bX+9)$ and setting equal to $X^4 - 2X^2 + 9$, gives no possible solutions. I.e. we get a+b = 0, and 27a+3b = 0, so a = b = 0, which does not work.

The difficulty of showing irreducibility gives us added reason to be glad we have the abstract dimension theory to work with.

V. In example III, the field $Q[\sqrt[5]{2}]$ only contained the one real root of X^5-2, but none of the 4 complex, non real, roots. If we look at the polyonmial X^3-2, it also has only one real root $\sqrt[3]{2}$, so if we adjoin it, we get a real field $Q[\sqrt[3]{2}]$ of dimension 3 over Q, but in which none of the non real roots of this polynomial are found. Now the quotient of two

cube roots of 2 is a cube root of 1, and conversely, the product of a cube root of 1 and a cube root of 2 is another cube root of 2, so the field containing all cube roots of 2, equals the field containing one cube root of 2 and all cube roots of 1. So if μ is the non real cube root of 1 with angle 120^0, i.e. $\mu = \cos(120^0) + i \sin(120^0)$, then $\mu^3 = 1$, and μ satisfies the polynomial X^2+X+1. This polynopmial is quadratic hence is irreducible over any field it does not have a root, hence over both Q and $Q[\sqrt[3]{2}]$. Thus the field $Q[\sqrt[3]{2}][\mu] = Q[\sqrt[3]{2},\mu]$ has degree 2 over $Q[\sqrt[3]{2}]$, which ahs degree 3 over Q. Thus $Q[\sqrt[3]{2},\mu]$ has degree 6 over Q, and has as Q basis the elements $\{1, \sqrt[3]{2}, \sqrt[3]{4}, \mu\sqrt[3]{2}, \mu\sqrt[3]{4}\}$.

I did not try to cook up inverse formulas in this field, but after much playing around think I got a single generator for this field, namely the expected element $\sqrt[3]{4}+\mu$, which seems to satisfy the polynomial
$X^6-9X^3-18 = 0$, but i did not check this. If it does satisfy it, then the polynomial must be irreducible, if I was right that the element generates the whole field.

VI. If we want the smallest field containing not just one, but all 4 roots of X^4-2, it is $Q[\sqrt[4]{2},i]$, of degree 8 over Q, with basis
$\{1, \sqrt[4]{2}, \sqrt[2]{2}, \sqrt[4]{8}, i, i\sqrt[4]{2}, i\sqrt[2]{2}, i\sqrt[4]{8}\}$. Once we adjoin the element $\mu = \sqrt[4]{2}$, the polynomial X^4-2 factors as $(X^2+\sqrt[2]{2})(X^2-\sqrt[2]{2})$

$= (X^2+\sqrt[2]{2})(X-\sqrt[4]{2})(X+\sqrt[4]{2})$. Thus adjoining a root of $(X^2+\sqrt[2]{2})$, such as $i\sqrt[4]{2}$ completes the extension. But we could just as well have adjoined simply i, a root of X^2+1, as we did.

I did not compute roots, but could so so if I knew a single generator, and the irreducible polynomial it satisifed. I guessed that $\mu = \sqrt[4]{2}+i$ should generate, and tested it by asking whether the five elements $1,\mu,\mu^2,\mu^3,\mu^4$ were all independent. (If they were then the polynomial satisfied by μ would have degree > 4, and since it also divides 8 it would be 8, so the element would generate. So I wrote out the elements in terms of the basis above, and got the following coefficient vectors
(1,0,0,0,0,0,0,0)
(0,1,0,0,1,0,0,0)
(-1,0,1,0,0,2,0,0)
(0,-3,0,1,-1,0,3,0)
(3,0,-6,0,0,-4,0,4)

Row reducing those by Gaussian elimination, gave the matrix
(1,0,0,0,0,0,0,0)
(0,1,0,0,1,0,0,0)
(0,0,1,0,0,2,0,0)
(0,0,0,1,2,0,3,0)
(0,0,0,0,0,8,0,4)

Since all rows have a pivot, these rows are independent, so those elements $1,\mu,\mu^2,\mu^3,\mu^4$ are independent, so the element μ does have degree 8 over Q, but I did not find a degree 8 polynomial it satisfies, so still do not know how to find inverses.

VII. If we adjoin a root μ of the irreducible polynomial X^3-X-1, then we get a field extension of degree 3 over Q, $Q[\mu]$, with basis $1,\mu,\mu^2$ where $\mu^3 = \mu+1$. Then we can multiply by just taking elements of form $a+b\mu+c\mu^2$, multiplkying them in the usual way, and when we get powers higher than 2, just use the rule $\mu^3 = \mu+1$, and hence $\mu^4 = \mu^2+\mu$, to lower them back down to at most degree 2, and write the product in terms of the basis $1,\mu,\mu^2$. Since we have the irreducible polynomial satisfied by μ, namely X^3-X-1, we can use Euclidean algortihm to find inverses.

Some easy ones we can find directly are as follows: since $\mu^3-\mu - 1= 0$, we get $\mu^3-\mu = 1$, so $\mu(\mu^2-1) = 1$, i.e. $\mu^{-1} = (\mu^2-1)$. Since then $\mu^2(\mu-\mu^{-1}) = 1$, then $(\mu^2)^{-1} = (\mu^{-1})^2 = (\mu-\mu^{-1}) = \mu - (\mu^2-1) = 1+\mu-\mu^2$.

As for $(1+\mu)^{-1}$, we get $(1-\mu)(1+\mu) = 1-\mu^2 = -\mu^{-1}$, so $-\mu(1-\mu)(1+\mu) = \mu\mu^{-1} = 1$, so $(1+\mu)^{-1} = -\mu(1-\mu) = \mu^2-\mu$.

That's all the examples we have time for now. If you are like me, these tedious computations make me long for the simplicity of the abstract approach to the subject. But there is no real substitute for examples, since the theory is meaningless without them.

4000/6000 Day 28, Compass and straightedge constructions

The Greeks had two favorite approaches to numbers, ratios of integers, and lengths of geometrically constructible segments. They knew how to construct all rational numbers as lengths, starting froma given unit length, and they knew that some constructible numbers were not rational, such as sqrt(2), the length of the diagonal of a unit square. Thus they knew that constructible numbers formed a larger realm than merely rational numbers but they were not sure how much larger it was. For instance, given a cube, they could imagine another cube with twice the volume of the first. But they did not know how to construct the edge of such a cube, given the edge of the first, using only compass and straightedge. We will sketch how to prove this is in fact inpossible, using our knowledge of field theory and dimension theory. They also knew how to bisect angles, and another question they asked was whether every angle could be trisected using only compass and straightedge. The answer to this is also no, and we will prove that too. They asked as well whether one could "square the circle", i.e. given a circle, whether one could construct a square whose area was equal to that of the circle. This too is not possible, but the proof uses calculus and would take a relatively large amount of time and effort to present, which we do not have left.

Let us say what we mean by a "construction", a constructible point, and a constructible real number.

One assumes given a plane, as in geometry, and two points given in that plane. The distance between those points is taken as a unit length. Then one admits the following constructions "with compass and straightedge".

1) Given any two distinct points, we assume we can construct the unique straight line passing through both of them.

2) Given any two points, we assume one can construct the unique circle centered at the first point, and passing through the second.

Next we define what we mean by a "constructible point".
1) The original two points are considered constructed.

2) Given any two non parallel constructed lines, their intersection point is considered constructed.

3) Given any constructed line and any constructed circle, their intersection points are considered constructed.

4) Given any two constructed circles, their intersection ponts are considered constructed.

Next we define a "constructible real number".
Given any two constructed points, the distance betwen them, measured using the distance between the two original points as unit, is considered a constructed real numbers, as is minus it. A real number is constructible, iff it (or its negative) is the distance between two constructible points.

Using the previous definitions, we can prove the following facts.

Theorems:

1) Given any line and a point on or off it, we can construct the unique line perpendicular to the given line and passing through the given point. (We must use the existence of some other constructed points to do so.) By constructing two perpendicular lines, we can construct a line parallel to a given line, and passing through given point.

2) Given any line and a point on it, and any constructible number d, we can lay off starting from the given point, another point on the given line and having distance |d| from it. I.e. we can lay off on any line with a chosen origin point, every point with constructible distance from the origin.

3) Thus we can erect, starting from our original line, and its two points, a coordinate system of two perpendicular axes, on which all constructible lengths are laid off on both axes.

Thus given any point in the plane, using 1), we can mark off on these axes its "x and y" coordinates, and given any two constructible x and y coordinates, we can construct the corresponding point in the plane having these coordinates. Consequently, a point in the plane is constructible if and only if both its x and y coordinates, with reference to our original line, and its perpendicular, are constructible numbers.
(This follows immediately from the two previous results.)

4) The constructible numbers form a subfield K of all real numbers, i.e. they are closed under addition and subtraction, multiplication, and division (by non zero numbers), and thus all rational numbers are constructible. (This uses similar triangles.)

5) If a line passes through two distinct constructible points with coefficients in a subfield F of K, then there is an equation for that line with coefficients in F.
(If the two points are (a,b), and (c,d) an equation is either $x = a$, if $a = c$, or $(y-b)(c-a) = (x-a)(d-b)$, if not.

6) If two lines have equations with coefficents in a field F, with respect to our coordinate system, then their intersection has coordinates in F also.
(Solving two simultaneous linear equations is done, as we know from math 2500, using only the field operations of addition, subtraction, multiplication and division, on the coefficients.)

7) If a circle has a constructible radius, and is centered at a point with constructible coordinates, all these numbers lying in a subfield F of K, then there is an equation for the circle also having coefficients in F.
(If the center is (a,b) and radius is r, an equation is $(x-a)^2 + (y-b)^2 = r^2$.

8) If a circle and a line both have constructible coeficients lying in a subfield F of K, then their intersection has coefficients in an extension field of degree 1 or 2 over F. The same holds for the intersection for the intersection points of two circles with equations in F.
(For example if the line has an equation in x,y in which the coefficients of y is non zero, then we can solve for y and substitute the resulting expression in x in place of y in the eqaution for the circle, obtaining a quadratic equation in x for the x coordinate of the intersection points. Then y

is linear in x.)

Consequently,

9) Any point which is constructible starting from our original two points, has coefficients in a field of degree some power of 2 over Q. I.e. if μ is any constructible number, then the degree of $Q[\mu]$ over Q, is some power of 2. (This follows from what we have said, using prop. 1.5)

In particular, a constructible real number cannot satisfy a polynomial which is irreducible over Q of degree n, if n is divisible by a prime other than 2. Thus, we deduce the following important results.

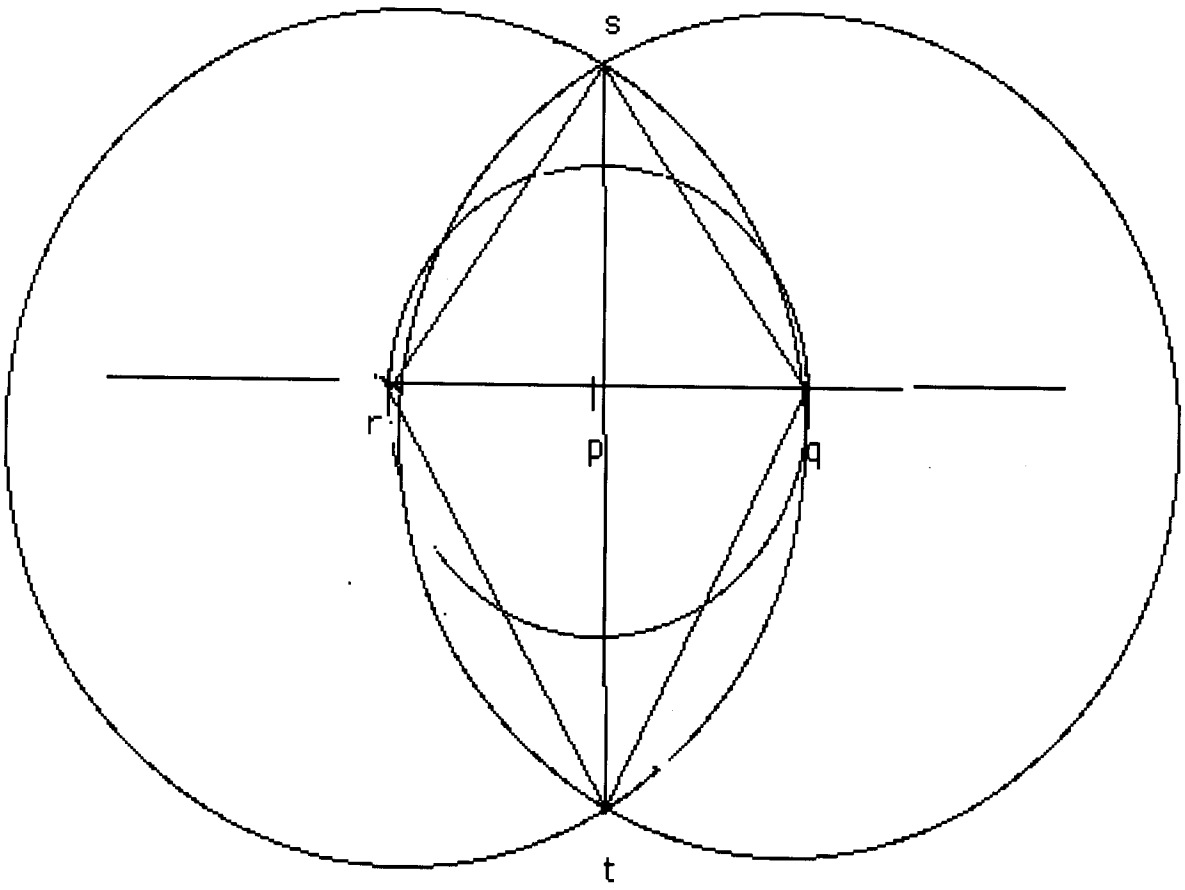
10) The real number cuberoot(2) is not constructible, nor is the real number $\cos(20^0)$, since we know each of these satisfies an irreducible polynomial of degree 3 over Q. Thus we cannot construct the edge of a cube whose volume equals 2, nor can we construct an angle of 20^0, using only compass and straightedge.

Remark: The deep result that the number π does not satisfy any polynomial with coefficients in Q (solved by Ferdinand Lindemann, building on important prior results of Charles Hermite), implies that we cannot square the circle.

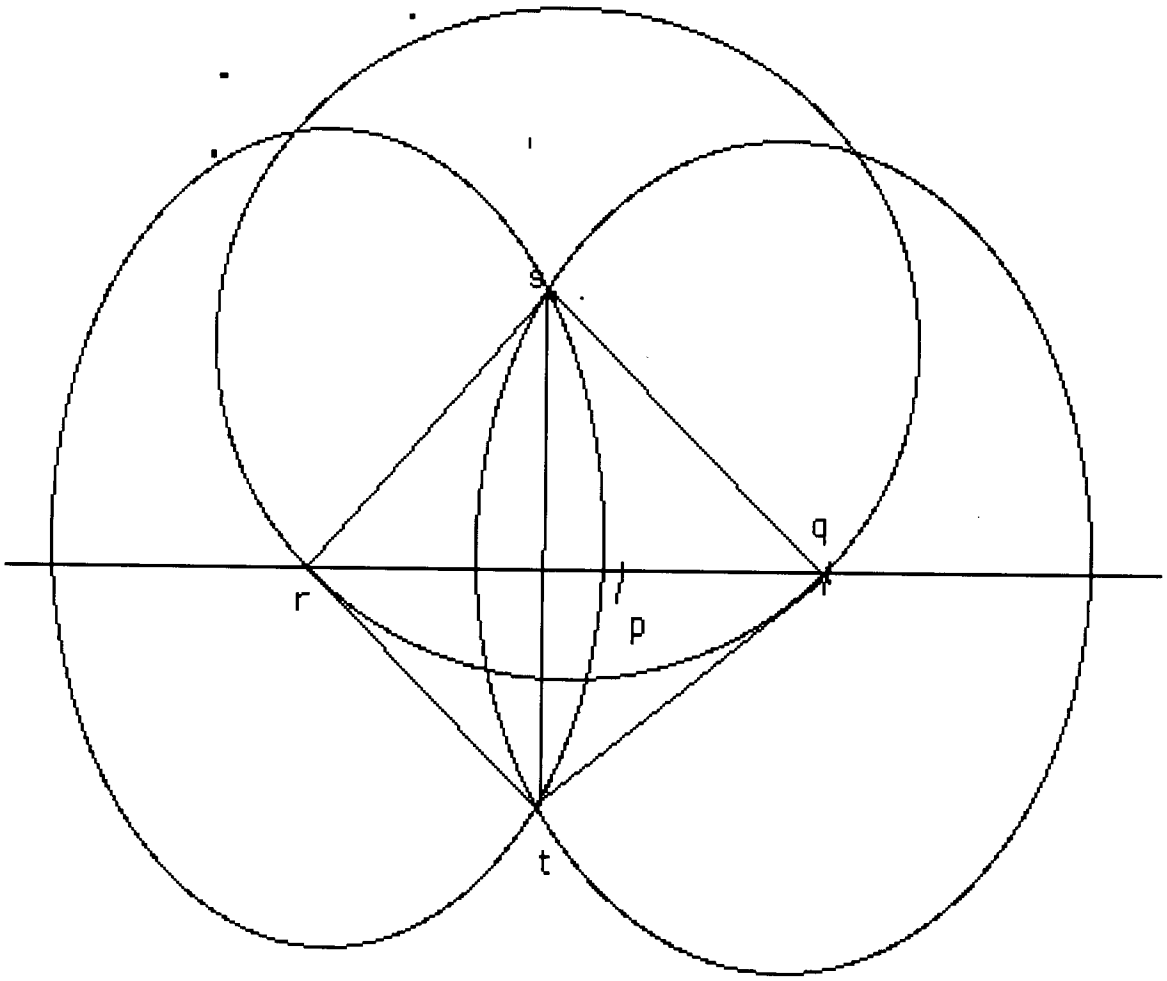
Proofs:

1) Given any line and a point on or off it, we can construct the unique line perpendicular to the given line and passing through the given point. (We must use the existence of some other constructed points to do so.) By constructing two perpendicular lines, we can construct a line parallel to a given line, and passing through given point.

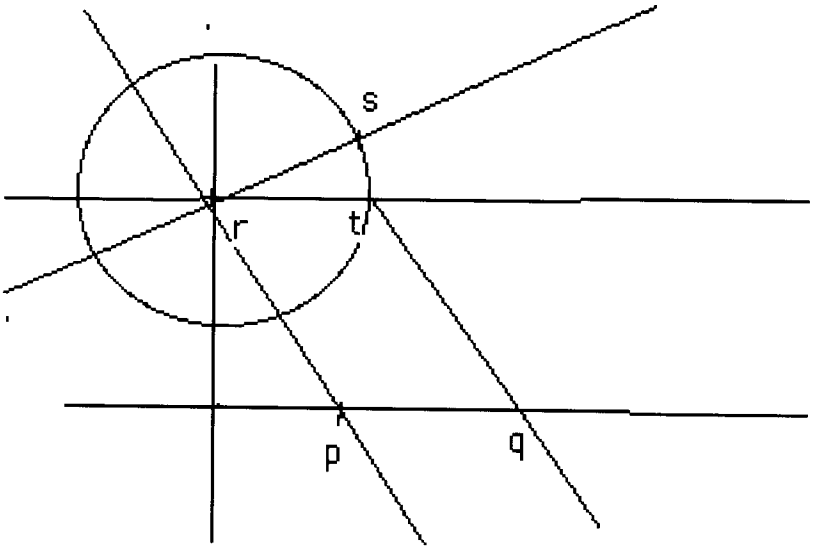
proof: Given a line and two points on it, p,q, construct the circle centered at p with radius |pq|, thus giving another point r on the other side of p, with |rp = |pq|. Then construct the circle centered at r with radius |qr|, and also the circle centered at q with radius |qr|. These two circles intersect above and below the line at points s,t. Then the quadrilateral q,s,r,t is a rhombus, so joining the two points s,t gives a line perpendicular to the original line pq, (since the diagonals of a rhombus are perpendicular bisectors of each other).



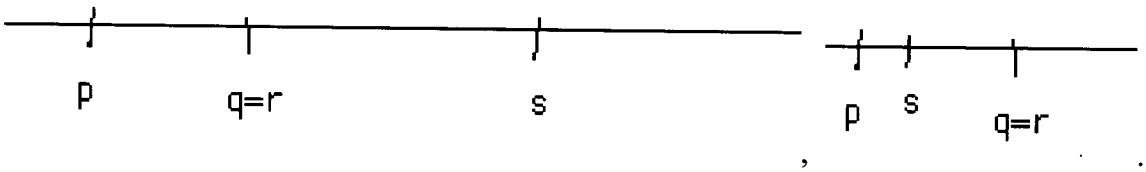
Next given a line, two points p,q on it, and a point s off it, it cannot true that both p and q are the unique closest point to r on the line. So the circle centeredt at s and passing through one of them, say q, is not tangent to the line, but meets it twice, at q and r. Then draw the two circles, centered at q and r, each passing through s. These two circles meet twice, at s and again at t on the other side of the line from s. Then the quadrilateral qsrt is a rhombus (all 4 sides have equal length), so the diagonals are perpendicular, hence the line st is perpendicular to the original line pq, and passes through s.



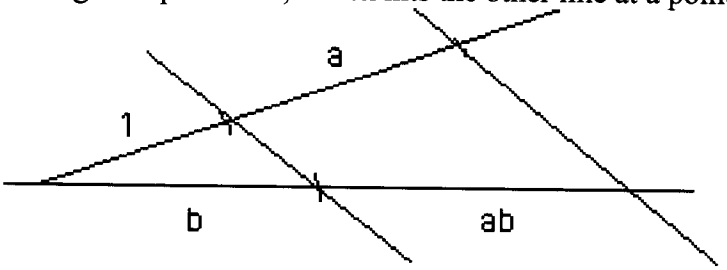
2) Given any line and a point on it, and any constructible number d, we can lay off starting from the given point, another point on the given line and having distance |d| from it.
proof: Let the line be given with point p on it, and let r,s be two constructible points anywhere. We draw a picture that covers most cases. If p = r, a circle centered at p of radius |rs| will do the job. If r lies on the line with p, we can erect a perpendicular to the line and mark off |rs| on that line twice, so that we may assume neither r nor s lies on the line. Then we construct the line through r parallel to the original line, and using a circle centered at r, with radius |rs| we construct t on that parallel line with |rs| = |rt|. Then connecting p to r, and constructing a parallel line to pr, passing through t, gives a line that meets our original one in a point q with |pq| = |rt| = |rs|.



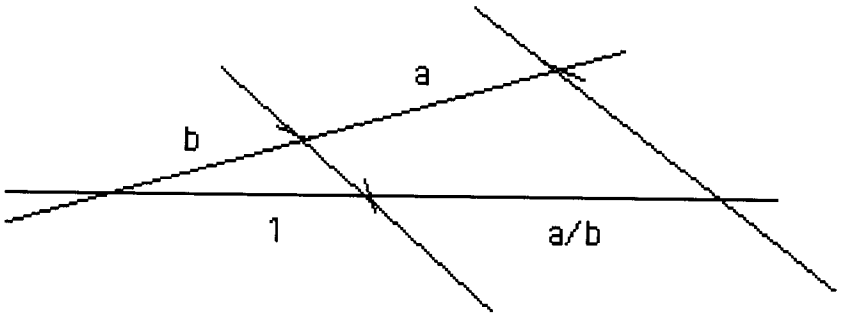
4) The constructible numbers form a subfield K of all real numbers, i.e. they are closed under addition and subtraction, multiplication, and division (by non zero numbers).
Use the previous construction to add |pq| to |rs|, by laying off the distance |rs| on the line pq, starting at q and going away from p. To subtract, go toward p.



To multiply a by b, use similar triangles constructed from parallels. I.e. given a,b, and a line with a point on it, construct two transverse lines through the point with the distance b laid off on one, and the the distances 1 and a on the other. Then join the points 1 and b, and construct a parallel through the point 1+a, which hits the other line at a point at distance ab from point b.



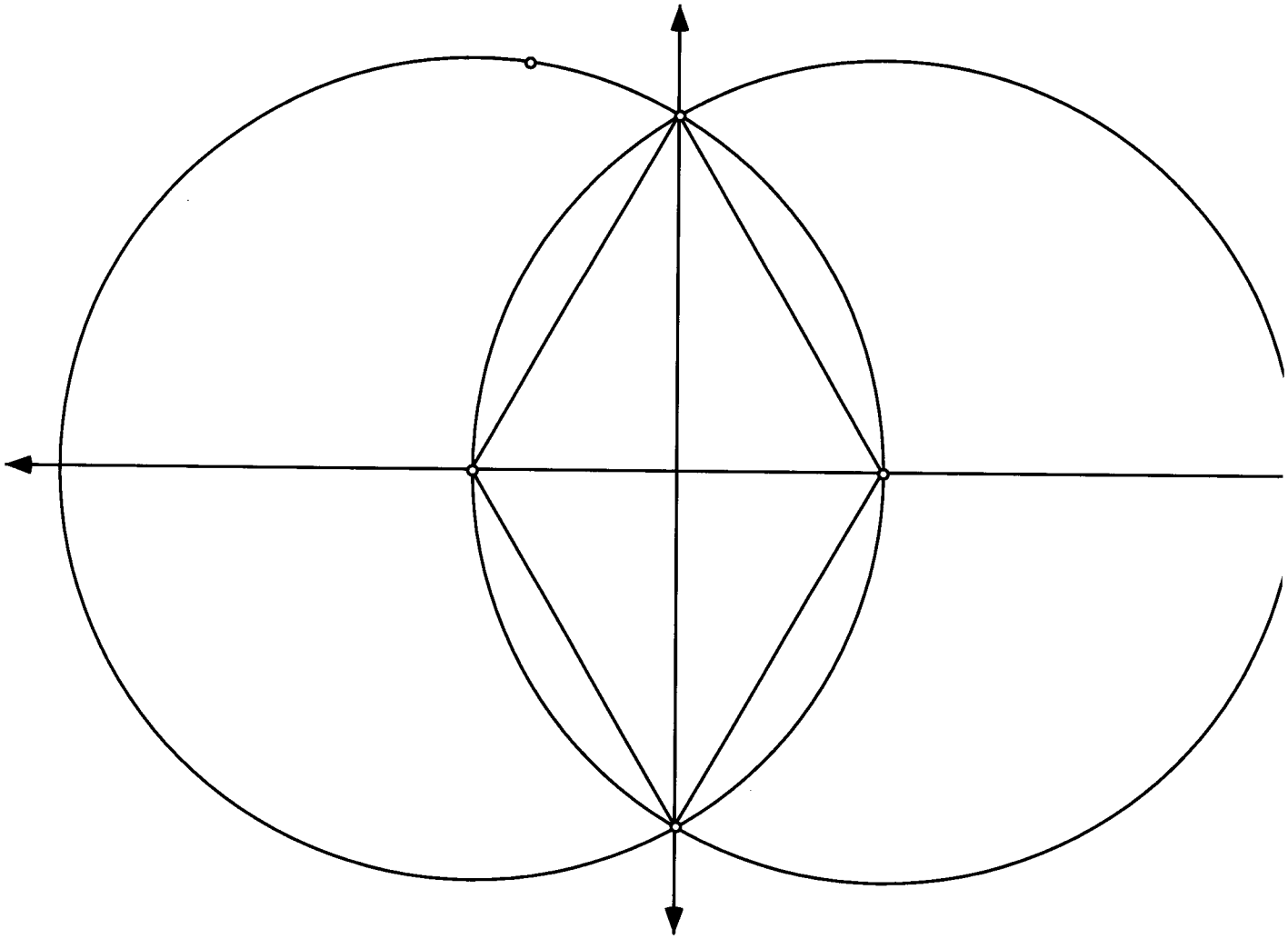
To divide, do something similar but in the other order.



I.e. given lengths b,a, laid off consecutively, lay off on another line through the origin, the length 1, and then connect b to 1, and draw a parallel through a+b, meeting the other line at a point a distance a/b from the point 1.

Since this is a subset of the real field, we do not need to check the commutativity, etc, all we need is closure under the field operations, so we are done.

Here is the final result of a geometer's sketchpad construction of a line segment perpendicular to a given line and bisecting the segment between two given points on the line.



Summary:
The main idea in 5.2 is that of a constructible real number, and that if μ is any constructible number then the extension $Q[\mu]$ has degree equal to some power of 2, over Q.

Start from 2 points in the plane and then construct all possible lines through any two constructed points, and all possible circles with a constructed point for center and pasing through another constructed point. For example, from just the original 2 points, there are exactly three possible constructions. Either the line through both points, or a circle centered at one of the points and passing through the other point.

Then consider the intersection of any two constructed figures, i.e. any two constructed circles, or lines, or line and circle, to be a new constructed point.

(For example if from the two original points we construct the one possible line, and the two possible circles, then we can intersect the line and the two circles and get two new points on the line. If we intersect the two circles we get two new points off the line such that joining them gives a line perpendicular to the original line.)

Then use the original two points as a unit distance, and with that unit, one can measure the distance between any two constructed points. Then a constructible real number is either the distance between two constructed points, or zero, or minus the distance between two constructed points.

Theorems: 1) The set of all constructible real numbers is a field containing Q (the rationals), and (unlike Q) also closed under taking square roots of positive elements.
2) The full set of constructible real numbers is infinite dimensional over Q, but if we choose any one constructible real number μ, then the field $Q[\mu]$ is finite dimensional over Q, and the dimension is always a power of 2.

The reason for the field structure is that one can add or subtract by laying off copies of a given length, and one can multiply and divide by using parallels and similar triangles. See the notes I gave out today for how to construct at least all rational numbers.

The reason for the dimension of $Q[\mu]$ being a power of 2, when μ is constructible, is that each constreuctible point is obtained by solving two simultaneous equations, either two linear equations, or a linear equation and a quadratic equation, or two quadratic equations.

Now solving two linear equations involves eliminating a variable and getting another linear equation, so the solution actually lies in the same field as the coefficients of the equations, i.e. the "extension" has degree 1. Solving one linear and one quadratic equation involves substituting the linear in the quadratic and getting a quadratic equation in one variable, which can be solved at worst by taking one square root using the quadratic formula. Thus the solution lies in a field extension of degree one or two over the field of coefficients of the two equations.

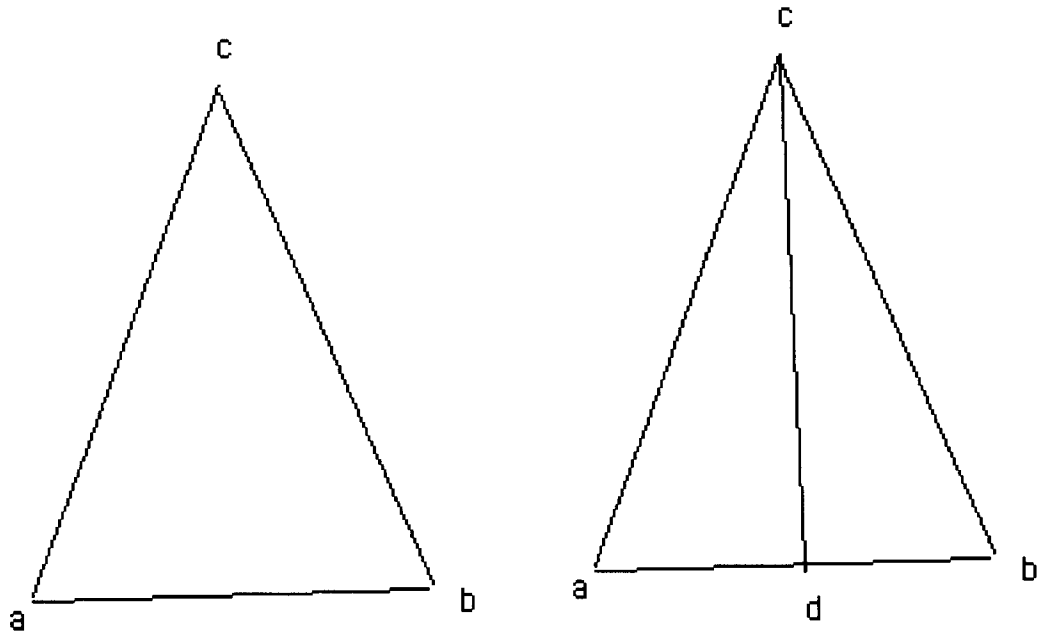
Solving two quadratic equations should involve getting a 4th degree equation, and thus one would think it could give an extension of degree 4. But the special thing here is that our two equations are both equations of circles, so they have exactly the same quadratic part, i.e. the only quadratic terms are x^2 and y^2. So if we subtract one of them from the other, the difference is a linear equation, and we are reduced to solving

one linear and one quadratic, as before. So again we get an extension of degree one or 2.

Thus every step in a construction involves a field extension of degree 2.
Thus doing any finite sequence of extensions, says (by prop 1.5, page 153, on multiplying degrees of field extensions), that the resulting point has coordinates in a field extension of Q whose degree is a product of 2's,
i.e. is a power of 2.

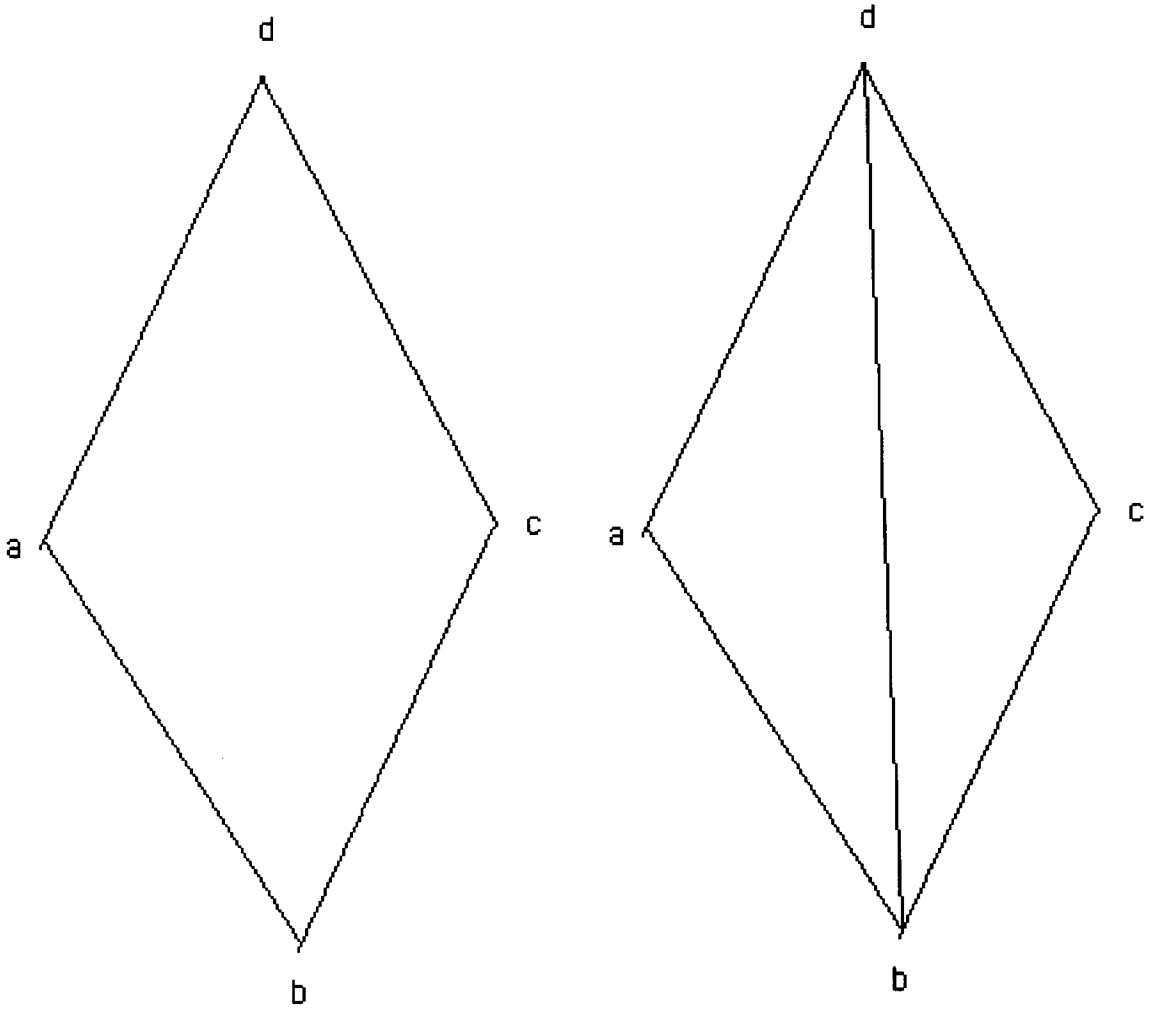
4000/6000 More compass and straightedge constructions
Assume that two triangles which have two sides of the same lengths, and with the angles between those sides also of the same measure, in fact have all sides of the same lengths as well as all angles. I.e. triangles are congruent if they share "SAS" or "side - angle - side". Also triangles are congruent by "SSS" or "side-side-side". I.e. if they have the same length sides then their corresponding angles also have the same measures.

Theorem: a triangle with two equal sides also has equal angles opposite those sides. (This is called an isosceles triangle.)
Proof: Look at this picture. Assume sides ac and bc are equal in length in the picture at eleft. Then drop a segment from c that bisects segment ab. I.e. assume in the following picture that ad equals db in length.

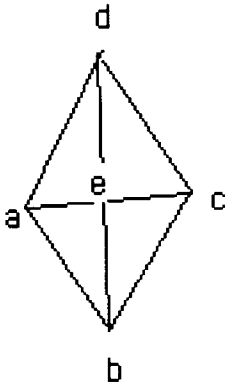


We claim triangles adc and bdc are congruent by SSS. I.e. they share a common side dc, and ac equals bc by hypothesis, and ad equals bd by construction. So they are congruent and hence have equal correpsonding angles. I.e. the angles opposite two equal sides are equal. So the angles opposite side dc are equal, i.e. angle a equals angle b, as claimed. **QED.**

Corollary: A quadrilateral with all 4 sides of equal length (rhombus), has diagonals which bisect each other and are perpendicular.
Proof: Consider the following picture.



In rhombus abcd, draw the diagonal db. Then triangles abd and cbd are congruent by SSS, since they share side bd, and the other pairs of sides are equal by hypothesis, i.e. cd equals ad, and cb equals ab. Then angle adb equals angle cdb. Now draw the other diagonal ac. Then triangles ade and cde are congruent by SAS, so angles aed and ced are equal. Since they add to a straight angle, both equal 90degrees. **QED.**

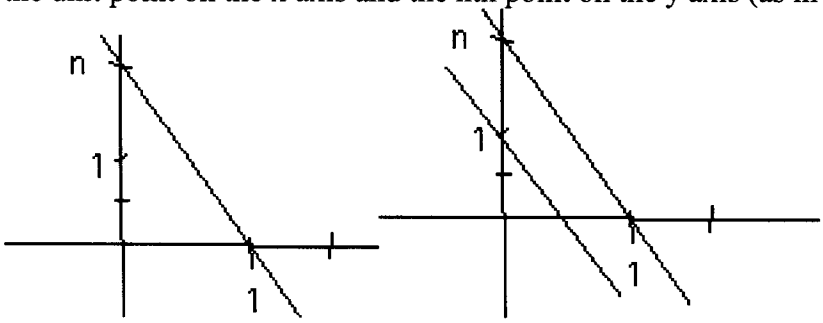


Corollary: It is possible to construct a line perpendicular to a given line and pasing through a

given point on or off the given line, (using the existence of two points on the line).
Proof: See yesterday's notes. **QED.**

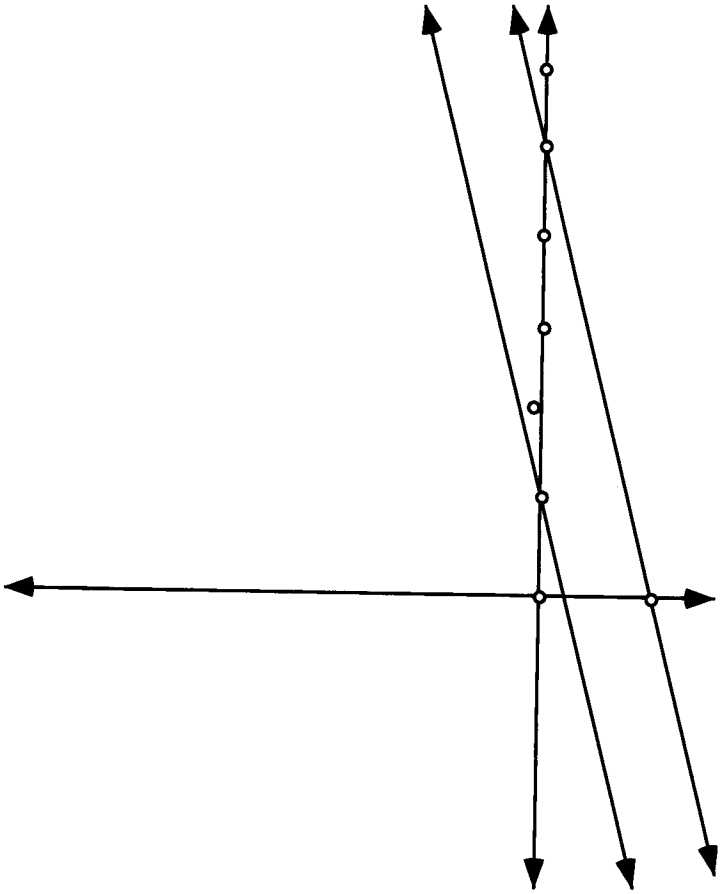
Corollary: Given a line L (and two points on it) and point p off the line, one can construct a line passing through p and parallel to L.
Proof: Construct a line M through p perpendicular to L as above. Then construct a line N through p perpendicular to M. Then N is parallel to L. **QED.**
Recall the principle of "similar triangles". If two triangles have equal angles, then any pair of sides opposite two equal angles have lengths in the same ratio. E.g. each side of one triangle has length equal to π times the length of the corresponding side of the other triangle.

Theorem: Given two distinct points on a line, one to use as origin and one as unit point, it is possible to construct on that line all points at rational distance from the origin.
Proof: integer points.
To construct the point 2, make a circle centered at the unit point passing through the origin. This circle meets the line, besides the origin, at point 2. To construct point 3, make a circle centered at point 2 and passing through the point 1. This circle meets the line besides at point 1, also at point 3. Continue to obtain any positive integer point, or do it in the other direction to get any negative integer point.
points at distance 1/n from the origin.
Erect a line perpendicular to the original line, passing through the origin (erect a "y axis"). The unit circle centered at the origin marks off a unit point on this new axis. Using this unit point construct all points on the y axis at integer distances from the origin. Now draw the line between the unit point on the x axis and the nth point on the y axis (as in left picture).



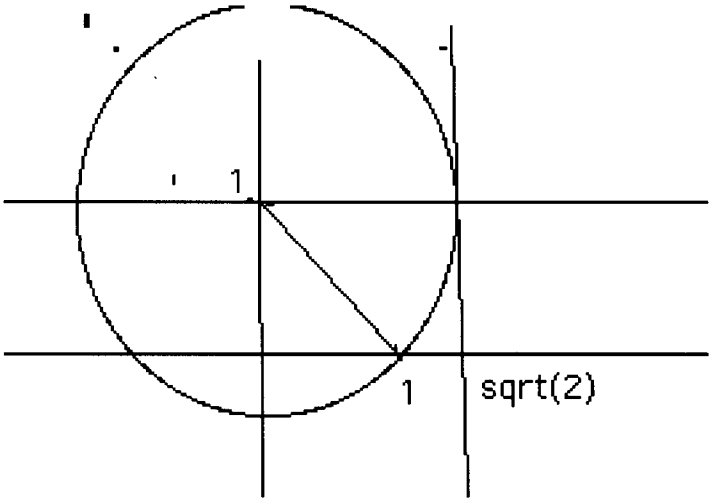
Then draw a line parallel to that line, but through point 1 on y axis. It will hit the x axis at the point at distance 1/n from the origin, by similar triangles. **QED.**
points at distance m/n from the origin.
Now imitate the same construction as done to find all integer points, i.e. the circle centered at point 1/n passing through the origin, meets the x axis further at point 2/n. In this way we obtain all points of form m/n on the x axis. Doing this for all integers m and $n \neq 0$, we obtain all rational points. **QED.**

Now I'll try this with geometers sketchpad. I tried to construct the point at distance 1/5 from my origin. But I did not like how slow it was to do all the constructions so I just guessed my unit distances. I also did not know how to label my points as "1", "2", ..."5". Anyway here it is.



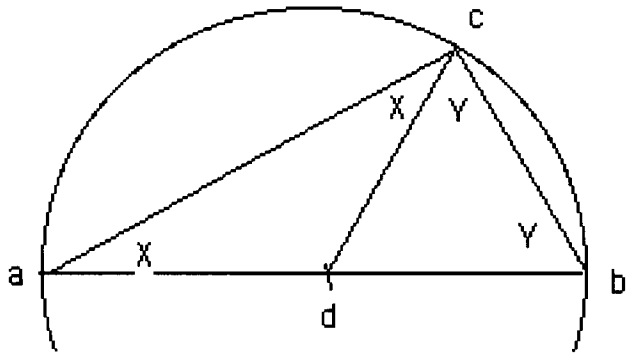
I guess I am not too good with this. This isn't much better than what I did just with macpaint.

Now I want to remind you that you can construct more than just rational points. For example you can construct any multiple of sqrt(2), or any number of form sqrt(2)/n, for any integer $n \neq 0$, as follows.
Make a circle centered at the unit point on the y axis, and passing through the unit point on the x axis, hence with radius sqrt(2). Make a line through the unit point on y axis, and parallel to the x axis, hence meeting the previous circle at distance sqrt(2) from the unit point on y axis. Then drop a perpendicular on the x axis from this point, hence meeting the x axis at distance sqrt(2) from the origin. Then imitate the earlier constructions of integer distances, i.e. of laying off copies of this given length, and also of dividing this length up into n equal parts.



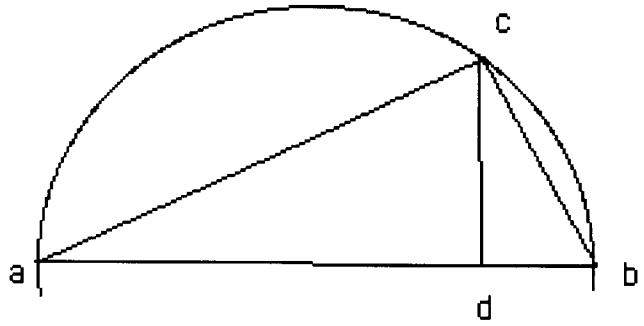
We can also construct the square root of any distance we already have, e.g. of any rational number, or repeated square roots of any rational number, e.g. 4th root of 2 as follows.

Lemma: Any triangle formed by joining the opposite ends a,b of a diameter of a circle, to any other point c on the circle has a 90degree angle at the third point c.
Proof: (See picture below.)



Draw a segment from c to the center of the circle at d. Then triangles adc, and bdc are isosceles, since each has two radii for sides. Thus angles dac and dca are both equal say to X, and angles dcb and dbc are both equal say to Y. Looking at the big triangle abc, its angles add to 180degrees, so 2X + 2Y = 180. Hence X+Y equals 90degrees, but X+Y equals angle acb. **QED.**

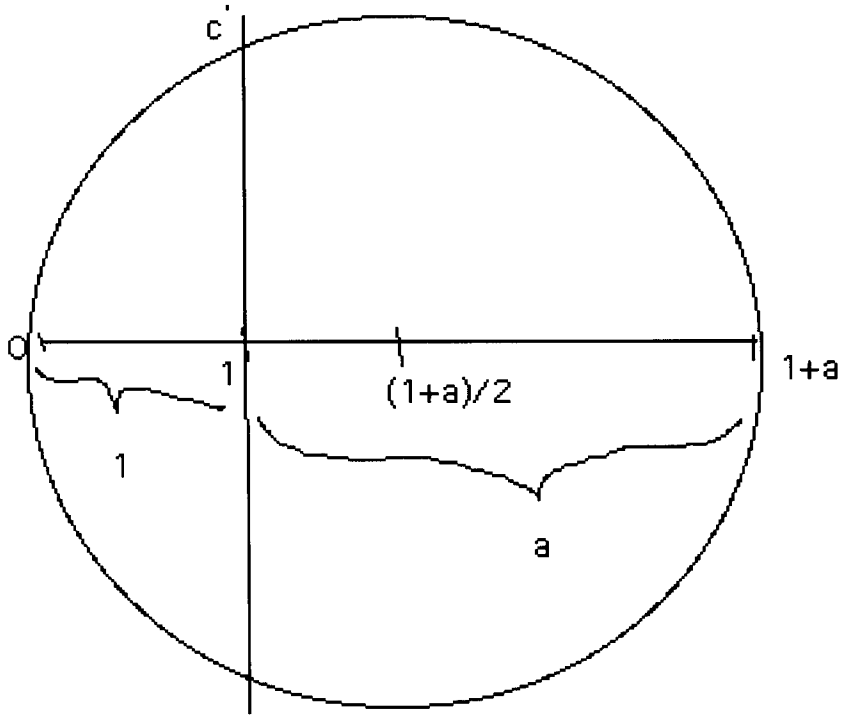
Lemma: Given any right triangle abc with 90degree angle at c, if we drop a perpendicular from vertex c to the hypotenuse, meeting it at d, then the product of the lengths ad and db equals the square of the height cd.
Proof:



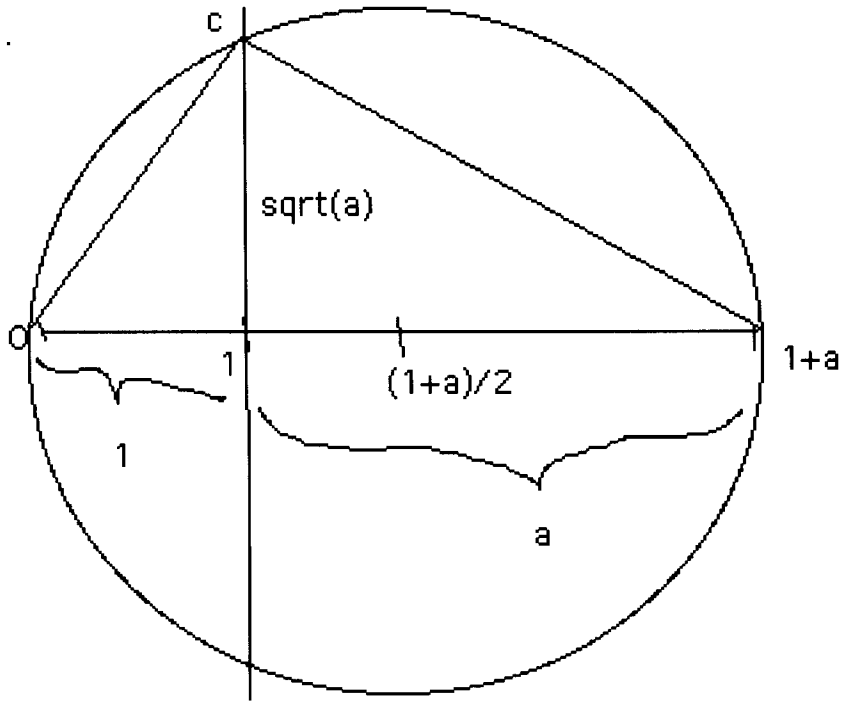
Since angles acd and bcd add to 90, as do angles acd and dac, then angle dac equals angle bcd. Hence also angles dca and dbc are equal, so the triangles adc and cdb are similar. Thus their sides are proportional, i.e. the ratios of the lengths |ad|/|dc| and |dc|/|db| are equal. Thus |ad| |db| = $|dc|^2$. **QED.**

Now let a be any given length. Then lay off length a starting from the unit point.

Then bisect this total segment, i.e. bisect the segment from 0 to 1+a, at (1+a)/2, and draw a semicircle centered at (1+a)/2, and with the segment as diameter. I.e. the circle is centered at (1+a)/2 and passes through the origin and through 1+a. Then erect a perpendicular to the segment, passing through the unit point and meeting the circle at the point c, as in the picture below.



Now draw the segments from 0 to c, and from c to (1+a), thus forming a right triangle as below. Then the height of this triangle (already constructed) has length equal to the square root of a, by the previous lemma.



By repeating this construction we can construct a segment whose length is equal to any real

number which lies in any real field extension of Q, obtained by repeatedly making degree 2 extensions. I.e. any real degree 2 field extension is obtained by adjoining a root of an irreducible quadratic polynomial to the given field. And by the quadratic formula, the roots of such a polynomial are in the field extension obtained by adjoining a square root of "b^2-4ac". Since we can construct this square root, and the constructible numbers form a field, we can also construct the root of any degree two polynomial whose coefficients we can construct. Thus the constructible numbers not only form a field, but a field which is closed under taking square roots. This is a very large field, infinite dimensional over Q, but it does not contain any roots of irreducible polynomials of any degree over Q except powers of 2. Thus cos(20^0) and cuberoot(2) are not constructible.

Corollary: The ancient problems of "trisecting all angles" and of "doubling any cube" are impossible by ruler and compass.
proof: These construction would yield algebraic numbers whose irreducible poynomials have degree 3 over Q. But if a real number A is constructible in a finite number of steps, then A is contained in some field extension of Q obtained by adjoining a finite number of square roots. Such a field has degree 2^n for some n over Q. But if Q[A] is contained uin a field of degree 2^n, then the degree of Q[A] over Q divides 2^n. But if A is cubert(2) or cos(20^0), we know the degree of Q[A] over Q is 3. Thus these numbers cannopt be contained in any finite field extension of degree 2^n, hence they are not constructible numbers. Thus the corresponding geometric constructions cannot be carried out. **QED.**

Remark: It was proved in the 19th century that π does not satisfy any polynomial in Q[X[at all, i.e. π is not only not constructible, it is not even algebraic over Q. Hence one also cannot "square the circle", i.e. one cannot construct by ruler and compass a square with the same area as a given circle.

It is possible to construct regular polygons of 3,4, 6, and even 5 sides. I.e one can construct the sin and cos of the angles $2\pi/3$, $\pi/2$, $\pi/3$, and also $2\pi/5$. Thus one can separate the circle into 3,4 5 or 6 equal arcs. But one cannot construct an angle of $2\pi/7$, since according to a book on my shelf, (Herstein's Topics in Algebra), $2\cos(2\pi/7)$ satisfies the irreducible cubic equation $X^3+X^2-2X-1 = 0$. I myself have never constructed a regular pentagon that I can remember. I do not think I learned that in high school. Can you do that? It was Gauss, 2,000 years after the Greeks, who discovered exactly which regular polygons, i.e. which angles of form $2\pi/n$, could be constructed with compass and straightedge. Moireover he constructed a regular 17 - gon! a construction described in my book as "correctly believed for centuries to be very difficult, but incorrectly assumed to be impossible."