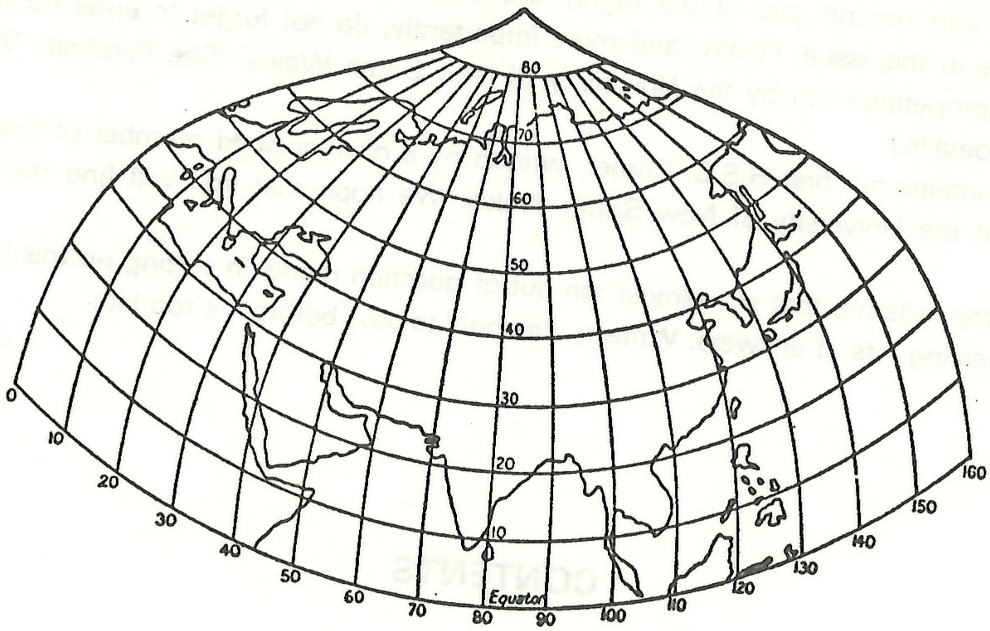
Parabola Volume 16, Issue 1 (1980)

MERCATOR, MAPS AND MATHEMATICS

John Mack*

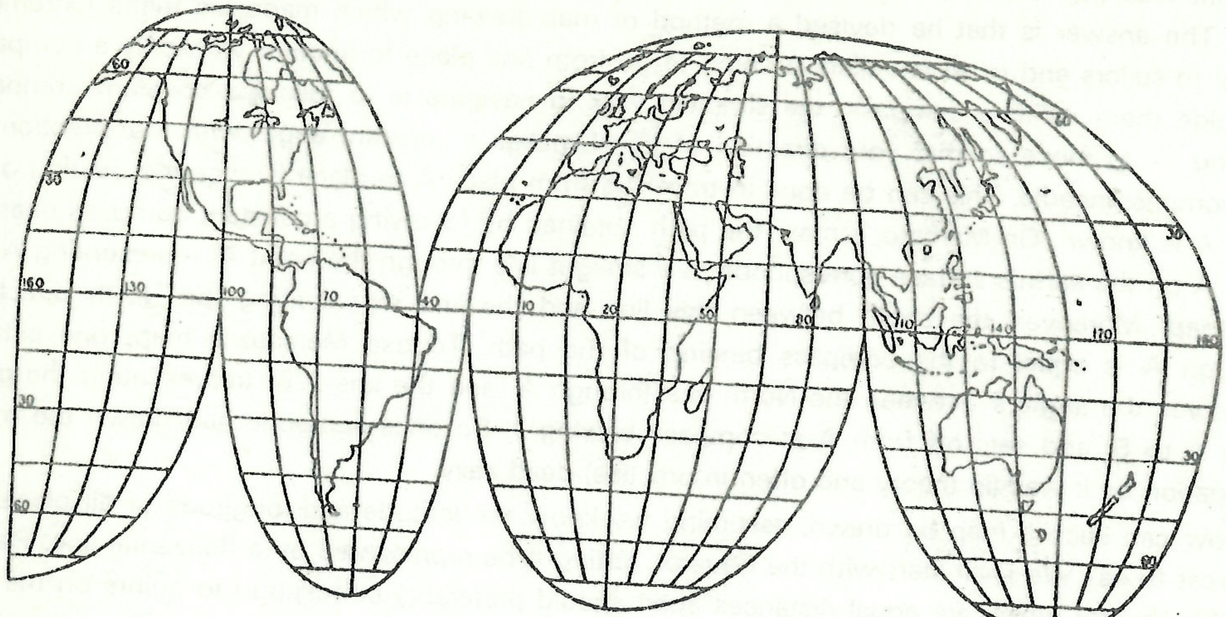
Newspaper readers may have noticed earlier this year, buried away on an inside page, a small item announcing that someone had recently found, in a Dutch bookshop under a pile of rubbish, an old book which happened to be the only known surviving copy of Gerandus Mercator's original World Map. This mid-16th Century map, bought for some $800 from the unknowing bookseller, is expected to sell for tens (or even hundreds) of thousands of dollars, and one imagines that this value is placed on it not merely because of its age and uniqueness, but also because of the fame attached to the discovery of a method of map-drawing that completely superseded previous methods and remains still in use today.



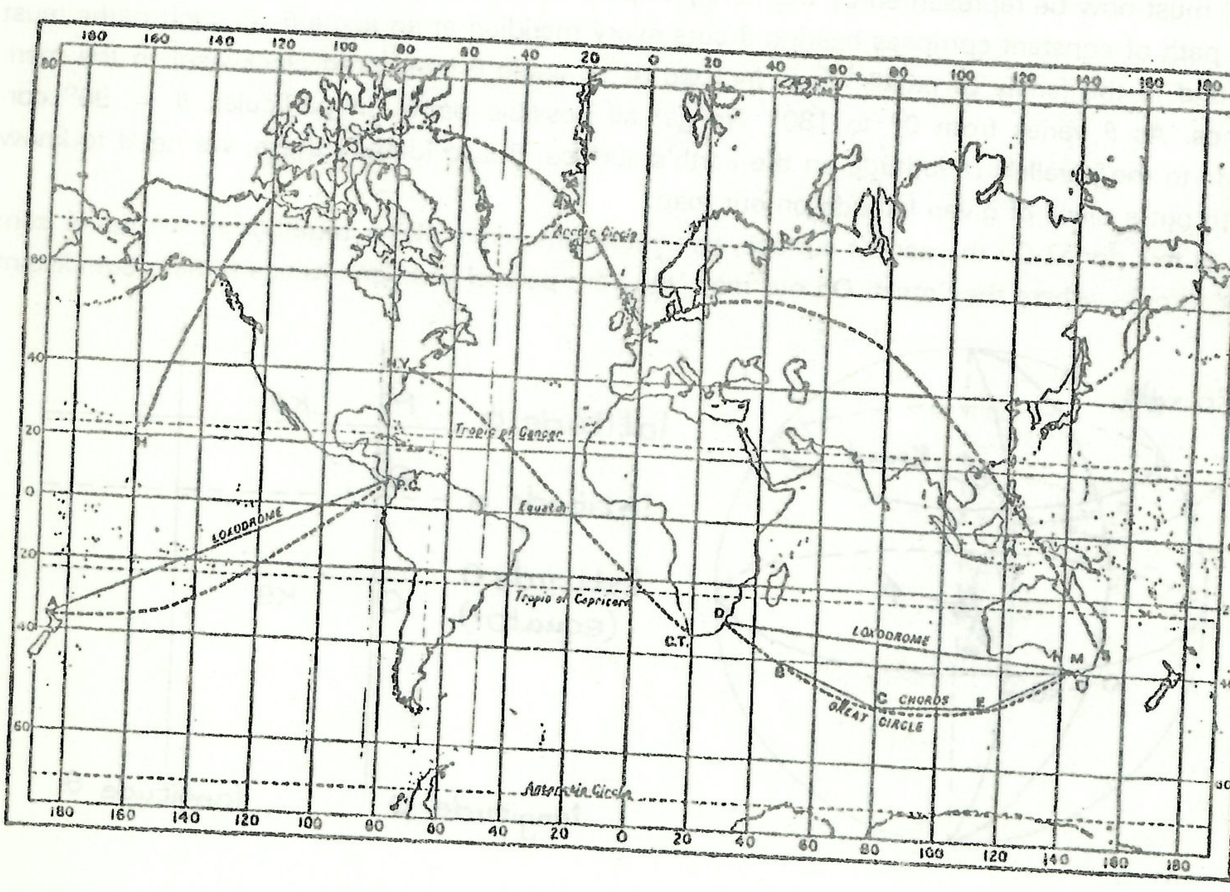
Eurasia on a Bonne's conical projection

If you have at home or at school a map of the world, it is quite likely that somewhere on it will appear the words "Mercator's Projection". Modern atlases still use Mercator's method of map-drawing for some maps, but a check of each atlas map will often produce a number of different methods — for example "conic projection", "Bonne's projection", "Lambert's Azimuthal Equal-area Projection", "Van Der Grinten Projection" and "Chamberlin Trimetric Projection" are some of the descriptions appearing on maps I have at home.

* *Dr. John Mack is a Senior Lecturer in the Department of Pure Mathematics at Sydney University.*



Interrupted Mollweide projection



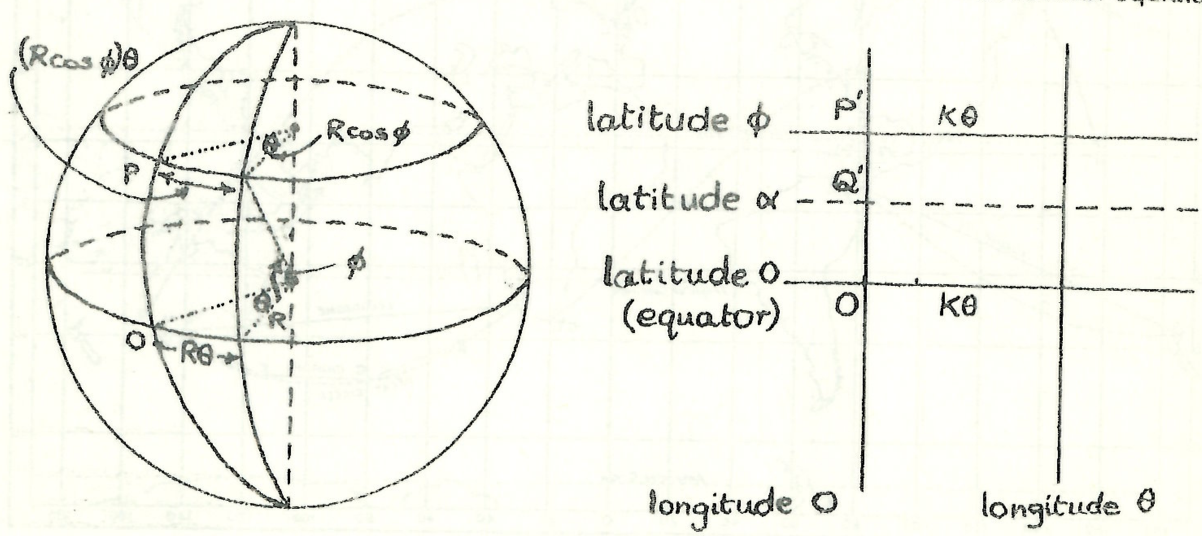
Mercator's Projection: ——————— rhumb line, – – – – – – – – great circle

What was the reason for Mercator's method superseding all others back in the sixteenth century? The answer is that he devised a method of map-drawing which made his maps extremely useful to sailors and to all travellers wishing to go from one place to another with only a compass to guide them. With a compass, the simplest way to navigate is to follow a constant compass bearing — to move so that your direction of travel makes a constant angle with the direction of the compass needle. This can be done in travelling from place A to place B once the bearing of B from A is known. On Mercator's map, the path obtained by following a constant compass bearing from A on the earth's surface corresponds to a straight line through the point A′ representing A on the map. Moreover, the angle between this line and the line representing the North direction through A is equal to the compass bearing of the path. To use Mercator's map, one simply measures the angle θ between the North line through A′ and the line A′B′ (representing the path from A to B) and sets off from A at compass bearing θ. So once someone had drawn the map, navigation by it was (in theory and often in practice) dead easy.

How can such a map be drawn, assuming we know the latitude and longitude of all places of interest to us? We shall start with the Equator, letting it be represented by a horizontal line. Places on the equator which are equal distances apart should preferably correspond to points on the line spaced evenly apart — so that one specification of scale will do for measuring all distances along the equator — so we choose a fixed distance to represent one degree of longitude at the equator.

The meridians (representing true North rather than magnetic North, a distinction we ignore at present) must now be represented by the family of parallel lines perpendicular to our equator. Also, since a path of constant compass bearing θ cuts every meridian at an angle θ, all such paths must correspond to the family of parallel lines inclined at an angle θ (measured clockwise) to the meridian lines. As θ varies from 0° to 180° we get all possible paths. In particular, $\theta = 90°$ corresponds to the parallels of latitude on the earth's surface. But to identify these, we need to know where to put a place of given latitude on our map.

How is this done? On the earth's surface, two meridians move closer together as we travel from equator to pole, where they meet. On our map, they are parallel lines and hence remain equidistant



The construction of Mercator's projection

from each other. This means that the scale we chose for equatorial distances is increasingly distorted as we approach the poles. At latitude ϕ, an easy calculation shows that the scale distance between our two meridians is a factor of sec ϕ too large — distances *along* the parallel are too big compared with the real distances corresponding to them on the earth's surface. But this distortion must apply to short distances in every direction about a point P on the parallel, because the angle preserving property of our map forces small triangles about P on the earth to correspond to similar triangles about the image of P on our map. Hence along the meridian at P, short scale distances are also a factor of sec ϕ too large. This applies to each point of a given meridian, and we can now work out where to place the point P′ on our map corresponding to P.

Draw the meridian line OP′ through P′, with O on the equator. We need to know the scale distance OP′ in terms of our equatorial scale units. At O, the distortion factor is sec O = 1. At Q′ (latitude α), the distortion factor is sec α. This means that the arc length on the meridian between latitudes α and $\alpha + \triangle\alpha$, which is just $R\triangle\alpha$, where R is the radius of the earth, must correspond to a scale distance $\triangle y = K \sec\alpha \, \triangle\alpha$ about Q′, where K is the equatorial scale distance corresponding to R. The approximate relation $\triangle y / \triangle\alpha = K \sec\alpha$ leads to the exact relation

$$dy/d\alpha = K \sec\alpha,$$

hence, integrating with respect to α from O to ϕ, we find

$$OP' = \int_0^{\phi} (dy/d\alpha)\, d\alpha = K \int_0^{\phi} \sec\alpha \, d\alpha.$$

This integral is fun to evaluate if you haven't seen it before. The neatest way to write the answer is

$$OP' = K \log \tan (\tfrac{1}{4}\pi + \tfrac{1}{2}\phi),$$

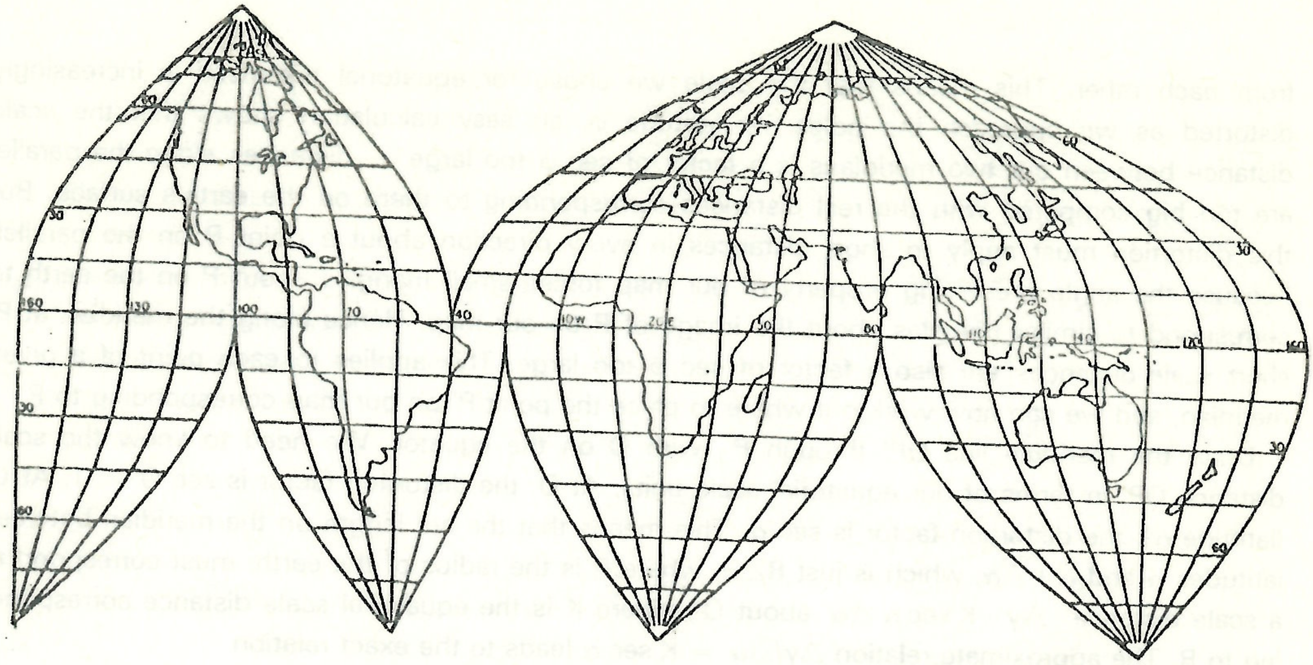
and we can now draw on our map a line for each parallel of latitude. Notice that we can never reach the North pole, and that our map will get bigger and bigger as latitude increases. (Greenland always does look much bigger than Australia!)

One of the properties of the map is that "small" areas preserve shape, so that the map is also a useful and recognisable representation of the surface features around us, wherever we are (except at the inaccessible poles). We could choose some other great circle (say 0° longitude) as our base line and then other areas of the earth become "inaccessible".

The paths of constant compass bearing on the earth's surface are called *rhumb lines* or *loxodromes*. They are not usually the shortest paths between two given points. Any two (non-diametrically opposite) points on the surface determine a unique great circle, and the shortest path between the two points is the smaller great circle arc joining them. This shortest path is clearly *not* a straight line on Mercator's map, and you might like to discover what curve corresponds to it.

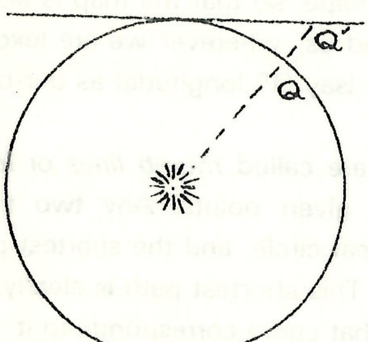
Another problem regarding Mercator's map is that of measuring distance along the rhumb lines. What is the formula, if any, for working out the real distance along the rhumb line joining the places corresponding to two given points P′ and Q′ on the map? How convenient is Mercator's map for measuring distances?

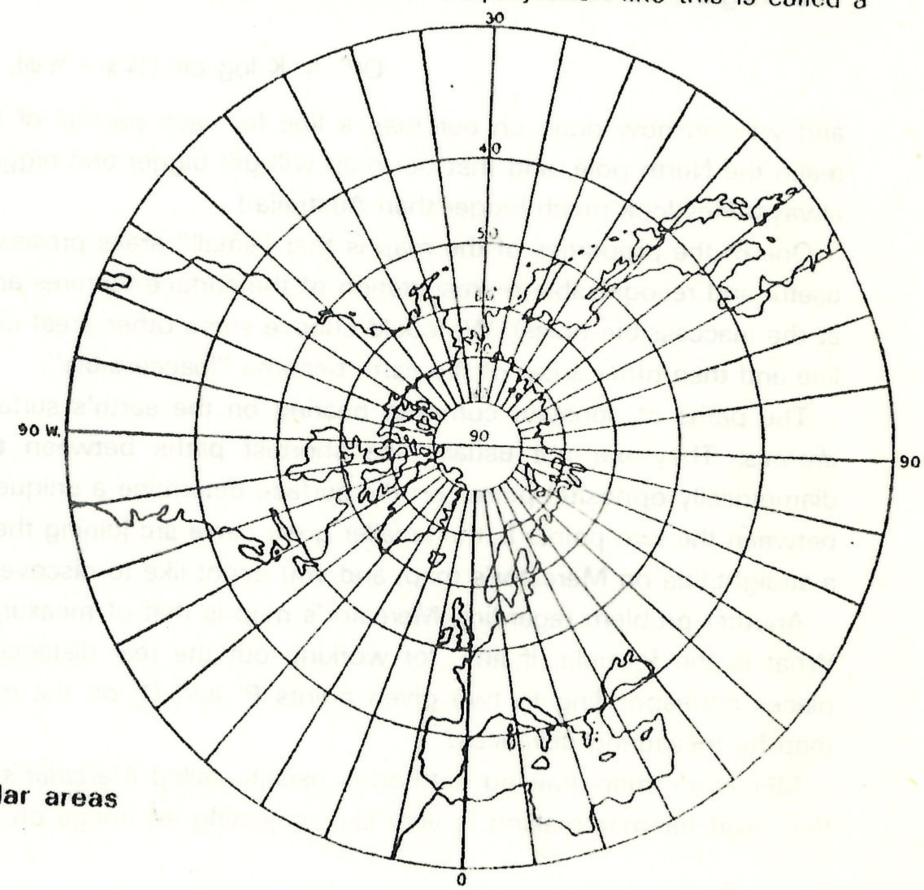
Mercator's map-drawing method is usually called *Mercator's projection*. Another type of projection used for map-making is very like projecting an image on a screen. Imagine a light source at



Interrupted Sanson-Flamsteed projection

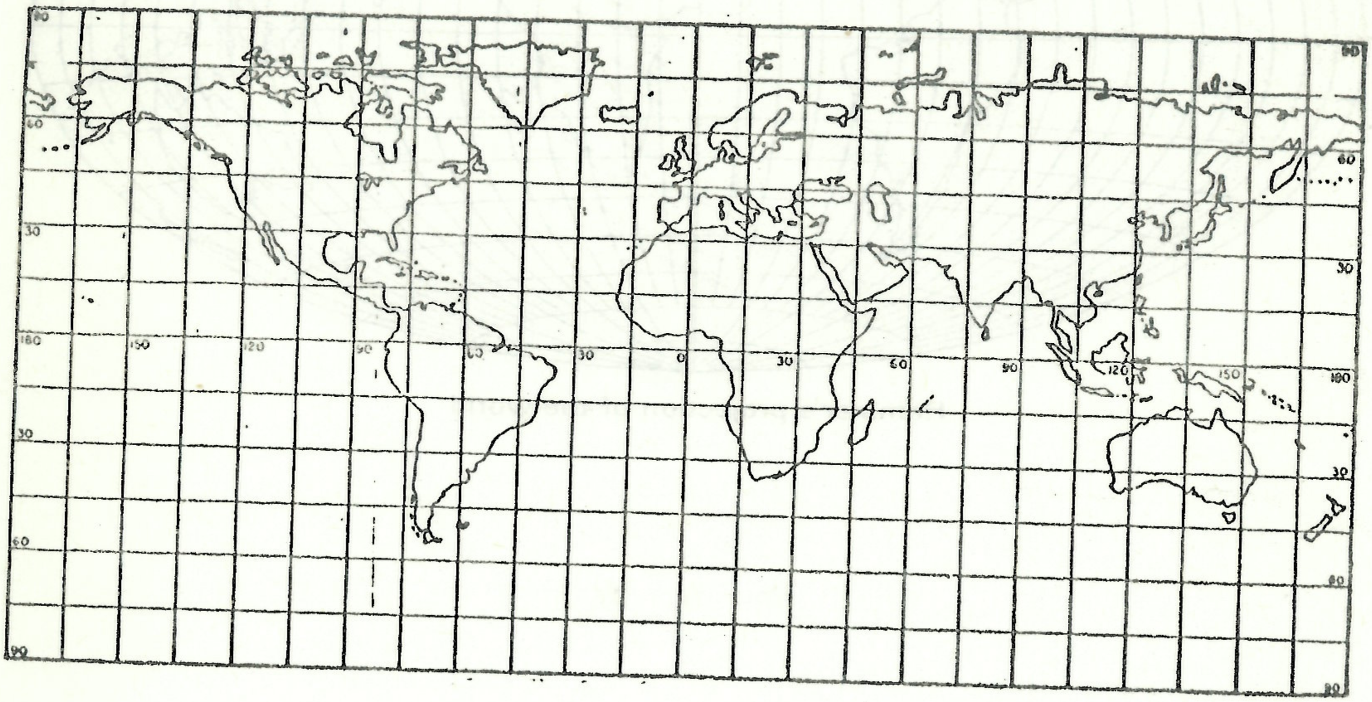
the centre of the earth (now suddenly transparent!) and the screen as a tangent plane touching the earth's surface at a selected point P say. Rays of light passing through a point Q on the surface strike the screen at a point Q', and in this way we obtain a map of the surface on the screen. This map is again a "nice" map for an area about P, but we expect some distortion as we move away from P. What properties would you expect such a map to have? For example, straight lines on the map correspond to what kinds of curves on the earth's surface? (A projection like this is called a gnomonic projection — why?)





Gnomonic map of north polar areas

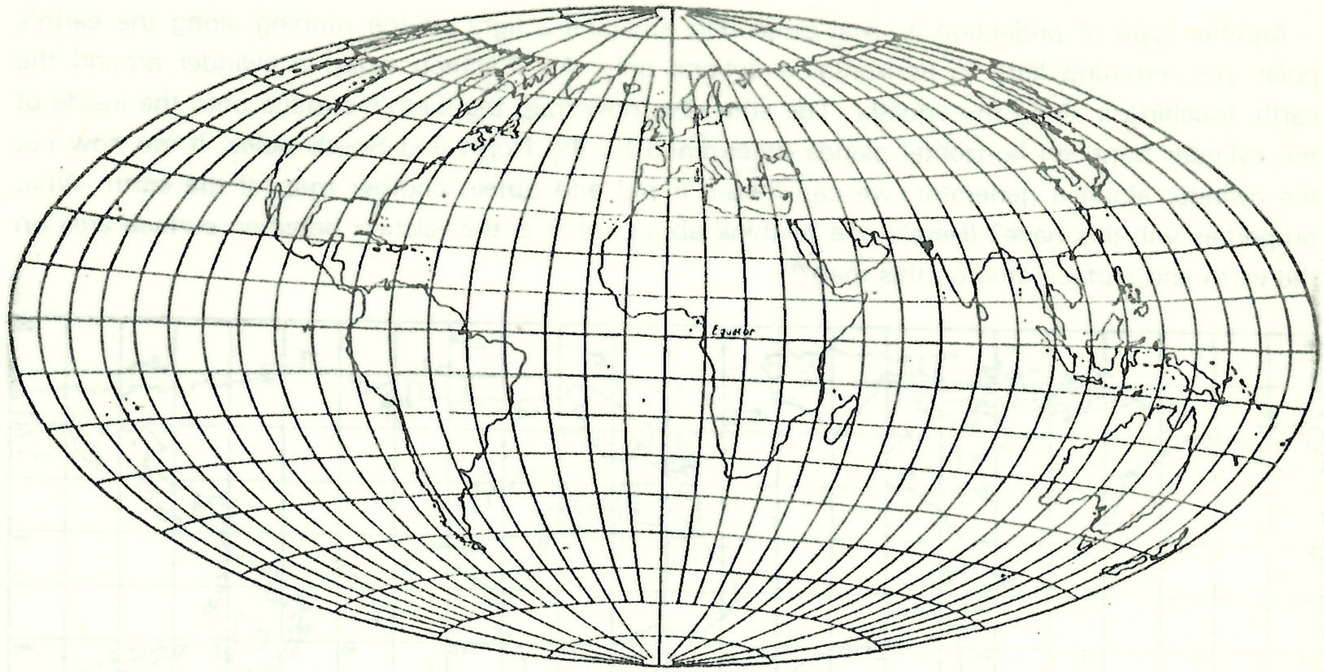
Another type of projection is axial projection. Imagine a light source running along the earth's polar axis, emitting light in horizontal directions only. As a screen, wrap a cylinder around the earth, touching it along the equator. The light rays now map the earth's surface onto the inside of the cylinder between horizontal planes determined by the North and South poles. If we now cut the cylinder along a generator, we can flatten it out and obtain another map of the earth. What properties will this have? (Here's one to think about: what is the relation between surface area on the earth and surface area on this map?)



Simple cylindrical map of the world (axial projection)

All maps of the earth that we use convey some kind of information about the earth that is interesting and useful. No one would be terribly interested in a world map that rendered the world unrecognisable. In other words, certain features of the earth, considered important to us, are expected to be preserved in some recognisable way in a map. Indeed, it often happens in mathematics that we need to represent one system (the earth) by a more convenient system (a plane map) so that the properties important for us are preserved. The best representation will depend on just what properties of the system we want to investigate.

To finish, return to world maps. On the surface of the earth, distance between points is measured as great circle distance. In the plane, distance is measured as ordinary euclidean straight line distance. Can we construct a plane map of a small area R of the earth's surface which is distance-preserving, that is it has the property that we can *always* measure the distance between any 2 points of R by measuring the ordinary distance between their image points on the map? Notice that Mercator's map certainly won't do. Any such map must map arcs of great circles onto straight line segments, and must also preserve angles. The extra condition we seek is that its scale remain constant over the entire map. Is such a map possible?



Hammer's projection of the world

☆ ☆ ☆

PRIME TIME

How fast can a computer run? Over recent years, considerable prestige has gone to the current holder of the record for the world's largest prime and it seems that each new idea in the computing field has cut its teeth on large prime numbers. As we explained in "The strange case of Father Mersenne" (Parabola, Volume 15, Number 2), the best way to find a large prime is to test the Mersenne numbers $M(n) = 2^n - 1$, by using an algorithm discovered by Lucas and Lehmer. The time taken to test the number M(n) on a given machine is roughly proportional to n^3. So the search for Mersenne primes provides quite a graphic measure of the development of computing power. The following table is a list of the known Mersenne primes and the times taken by their discoverers to check their primality.

	n	*M(n)*	*Discoverer*	*Year*	*Computing time*
1	2	3	Mersenne	1644	
2	3	7	"	"	
3	5	31	"	"	
4	7	127	"	"	
5	13	8191	"	"	
6	17	$2^{17}-1$	"	"	
7	19	$2^{19}-1$	Euler	1750	
8	31	$2^{31}-1$	"	"	
9	61	$2^{61}-1$	Lucas	1876	
10	89	$2^{89}-1$	"	"	
11	107	$2^{107}-1$	"	"	
12	127	$2^{127}-1$	"	"	
13	521	$2^{521}-1$	Lehmer	1952	1 minute
14	607	$2^{607}-1$	"	"	1 minute
15	1279	$2^{1279}-1$	"	"	13½ minutes
16	2203	$2^{2203}-1$	"	"	1 hour
17	2281	$2^{2281}-1$	"	"	1 hour
18	3217	$2^{3217}-1$	Riesel	1957	5½ hours
19	4253	$2^{4253}-1$	Hurwitz	1961	50 minutes
20	4423	$2^{4423}-1$	"	"	50 minutes
21	9689	$2^{9689}-1$	Gillies	1963	1½ hours
22	9941	$2^{9941}-1$	"	"	1½ hours
23	11213	$2^{11213}-1$	"	"	2¼ hours
24	19937	$2^{19937}-1$	Tuckerman	1971	35 minutes
25	21701	$2^{21701}-1$	Noll and Nickel	1978	8 hours
26	23209	$2^{23209}-1$	Noll	1979	8¾ hours
27	44497	$2^{44497}-1$	Nelson and Slowinski	1979	8 minutes

Lehmer's calculations were the first to use an electronic computer. His calculations were limited not so much by computing time but by the size of the machine. In fact, when the prime testing algorithm and the number to be tested were loaded into the machine, there was hardly any room left for the calculations. It took great ingenuity in those days to do the most elementary calculations and to prevent the computer from blowing valves right, left and centre. Let us compare the most recent prime M(44497) with Lehmer's prime M(1279). The index is about 35 times larger, so Lehmer's time of 13½ minutes for testing M(1279) extrapolates to 13½ × $(35)^3$ minutes (= 57 weeks) for M(44497), according to the rough rule of thumb mentioned earlier. So the latest computer is about 72000 times faster than the earliest machines. The prime M(44497) was discovered on a machine called the Cray-1 which employs rather refined parallel processing, so that it can perform many operations simultaneously. This makes it so complicated that only another computer is capable of feeding it. What are the prospects for 1984?

☆ ☆ ☆