THE SIGN FOR ZERO

Subhash C. Kak
Louisiana State University, Baton Rouge

The history of numerals holds a dual fascination since development of science presupposes numbers and the usage of numbers by the common folk can reveal surprising aspects of cultural life. Many of these dual issues for different civilizations are explored in Karl Menninger's *Number Words and Number Systems* (1958, 1969). But the story of how number systems evolved in early times is not very clear. Perhaps this evolution can be best seen in the Indian civilization, which since the Rigvedic age seems to have been interested in numeric questions. Georges Ifrah in his *From One To Zero: A Universal History of Numbers* (1981, 1985) has given an interesting account of the history of the Indian numerals. These are the common numerals and they are also called the Hindu-Arabic numerals. Ifrah has sketched a plausible explanation for how the place value system of the Indians may have arisen upon the use of the counting boards. The place value system with a clear use of zero goes back at least to +458 where it is used in *Lokavibhāga*, a Jain work on cosmology. The earliest epigraphical evidence relating to the use of the nine numerals in a place value manner goes back to +595 on a copperplate deed from Sankheda (Datta and Singh 1935). But some symbol for zero must have been in use in India as early as -200 since Piṅgala in his *Chandaḥśāstra* describes a representation system that is binary, as in modern computers, that would clearly need a symbol for zero. And much before Piṅgala, the grammarian Pānini used the notion of the null (zero) operator in his *Aṣṭādhyāyī*.

One question that has not been addressed by Ifrah or earlier historians is the shape of the zero sign. Why is it a circle or an oval rather than some other shape? The evolution of the shapes for the other signs is well understood and so it will not be recounted here, but it has generally been assumed

Reprinted from *The Mankind Quarterly* Volume 30, No. 3, Spring 1990

that the sign for zero appeared suddenly. The objective of this note is to provide an answer to this question and sketch how the final form must have evolved. We wish to sketch also a developmental process that is likely to have lead to the round form of the zero sign. This allows us to be more definite about the epoch when the sign was developed. We will also present arguments as to why the zero sign in India came to be fully developed with both the place value notion as well as the notion of null value associated with it.

That these questions are relevant is seen by noting that the Mayans and the Babylonians who, independently, invented their zeros picked very different forms. Furthermore the Babylonian and the Mayan zeros were not fully developed in their conceptual and operational meaning.

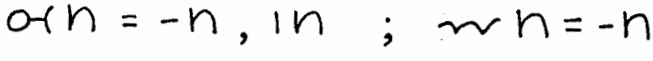
Numbers in Sanskrit have been expressed in different orders. Numerals were expressed in symbolic form in the place value notation in the order starting from the units. This occurs both in *Lokavibhāga* as well as *Āryabhaṭīya*. This order of representation is the order in which numbers are expressed in Classical Sanskrit. Thus 14 is *caturdaśa* (4+10), and 517 is *saptadaśādhikpañcaśatam* (7+10+500). While this is how numbers up to 99 are built up in Indo-Aryan languages, the numbers above 100 were actually represented in the reverse order excepting for the last two places. This is true also for Sanskrit as confirmed by checking numerical references in *Śulvasūtras* (Sen and Bag 1983). Thus in Baudhāyana SS., 225 is expressed as *dve śate pañcavimśatiśca* (BSS. 16.8) which is 200+5+20, or 187 is written as 100+80+7 (BSS. 11.2). Likewise *Vedāṅga Jyotiṣa* (Sastry 1985) expresses numbers sometimes in the same style. In other words there existed two styles for expressing numbers that were well established: one, in formal texts where to be consistent with the structure of numerals until 99 all numbers were expressed starting with the units; second, actual writing in terms of numeral signs was in the usual place value form starting with the highest order. This is confirmed by epigraphical evidence (Datta and Singh 1935) as well implied by rules for various mathematical operations in *Āryabhaṭīya, Brāhma-Sphuṭa-Siddhānta* and other

mathematical works.

Brahmi was in general use in India during the centuries that the sign for zero is likely to have been developed. Now 10 in Brahmi before the advent of zero was written as a fish sign,  . In later forms it was also written as with a single curving stroke (), or with the Asokan 1 attached to a circle ().

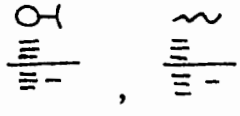
Using these forms one can now write numbers such as 15 as

where  is the Brahmi 5. In these forms, once the place value notation had become clear, we have the identities



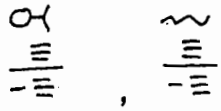
since 1 was written as a vertical stroke in Aśokan Brahmi. Therefore it is not unlikely that the shape of zero was determined by the oval related to the fish sign of the Brahmi 10. In numbers such as the one above, the zero sign clearly had the null (*śūnya*) value which explains its name. We also see how the two concepts expressed by the Indian zero, namely those of the place value and that of nothing, are likely to have become evident.

The significance of the reversed order usage may have played a role in the clarification of the concept of zero as 'null'. A sum such as 10+3 would now be performed as



and here as well zero would the value 'nothing'. In the

standard form:

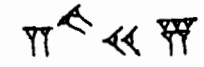


Perhaps the simultaneous existence of the two forms of expressing numbers helped in the development of the dual concepts associated with the zero sign.

We encounter the form for 10 in the 1st and 2nd century A.D. Nasik inscriptions, and in the 1st and 3rd century Andhra and Kṣatrapa inscriptions. And the curved form is seen in the 4th century Jaggayapeta and Pallava grants (Datta and Singh 1938). Therefore it is conceivable that the development of the zero sign occurred in these epochs.

But the above epochs do not provide a definite era for the discovery of the zero sign, since it is likely that the new usage competed with the traditional number system for centuries. In fact one would expect that inscriptions and deedplates would tend to follow the older and more commonly known style for a long time. For a parallel consider Europe where it took the Indian numerals more than 600 years to oust the less convenient Roman numerals. Even in India the older additive system with special signs for 10, 20, 30, and so on continued to be used, alongside the place value system, for centuries.

The development of the zero sign in India was motivated by numerical calculations. This is to be contrasted from the manner in which the zero signs arose in Babylon and Mexico, where the motivation was from the areas of astronomy and calendrical calculations. The Babylonian astronomical tablets use a sexagesimal numeration system. But it is imperfectly developed being partly additive and partly place valued as within the base of 60 a decimal system is used. The Babylonian system has only three specific symbols, namely those for 1, 10, and the later symbol for 0. Unless the groups of wedge marks are separated it is always possible to miscalculate the indicated number. This is seen in the example below:



This can be interpreted either as 2;0;25 = 7225 or as 2;0;20;5 = 217205.

The Mayans, on the other hand, used a vigesimal system but with a serious irregularity since its units were 1, 20, 18x20, $18x20^2$, $18x20^3$, and so on. Thus in this system the glyph representing a seashell (which is the 0) does not work as an operator, as it should in a true place value system. Furthermore the numbers up to 20 are additive as in the case of the Babylonian system, and therefore there exists the same possibility of ambiguity. The surviving inscriptions and codices do not write the numbers without specifying the units, which eliminates ambiguity but shows that the abstract nature of the place value number system was not fully understood. Clearly this system was also not designed for the needs of ordinary calculations. The rationale behind the Mayan system was the counting of the days of 18 months, each of 20 days.

In view of the demonstration of the derivation of the Brahmi numerals from the earlier Indus numerals (Kak 1989, Kak 1990), especially how the Indus 5, and 10 became the corresponding Brahmi signs, the above derivation pushes the ancestry of the zero sign to the third millennium B.C.

The choice of a circle, or sky, or space *kha* for zero had a happy reinforcement from the Brahmi sign for *kha* which is a circle with a hook on top of it.

To conclude, we see that the development of the concept of zero and its form in India was related to the form of the Brahmi sign for 10. Furthermore we see that the explication of the full power of the symbol may have been aided by the two different ways of expressing numbers in Sanskrit. We have identified a late form of Brahmi 10 as the immediate parent of the zero sign.

REFERENCES

Datta, B. and Singh, A.N.
1938 *History of Hindu Mathematics*. Bombay: Asia Publishing, 1962.

Ifrah, Georges
1985 *From One to Zero*. New York: Viking Penguin.
Kak, S.C.
1987 On the decipherment of the Indus script - a preliminary study of its connection with Brahmi. *Indian Journal of History of Science*. 22,51-62.
1988 A frequency analysis of the Indus script. *Cryptologia*. 12, 129-143.
1989 Indus writing. *Mankind Quarterly*. 30, 113-118.
1990 Indus and Brahmi: further connections. *Cryptologia*. 14. 169-183.
Menninger, Karl
1969 *Number Words and Number Symbols*. Cambridge: The MIT Press.
Sastry, T.S. Kuppanna
1985 *Vedāṅga Jyotiṣa of Lagadha*. New Delhi: Indian National Science Academy.
Sen, S.N. and Bag, A.K.
1983 *The Śulbasūtras*. New Delhi: Indian National Science Academy.