Tārīkh-e ʿElm: Iranian Journal for the History of Science, 6 (2008), pp. 37-48

A Brief History of Zero

George Gheverghese Joseph
University of Manchester george.joseph@manchester.ac.uk

(received: 01/06/2008 - accepted: 15/10/2008)

Abstract

It is generally recognized that 'zero' as we understand the concept today originated in two geographically separated cultures: the Maya and Indian. However, if zero merely signified a magnitude or a direction separator, the Egyptian zero, ***nfr***, dating back at least four thousand years, amply served these purposes. If zero was merely a place-holder symbol, then such a zero was present in the Babylonian positional number system before the first recorded occurrence of the Indian zero. If zero was represented by just an empty space within a well-defined positional number system, such a zero was present in Chinese mathematics a few centuries before the beginning of the Common Era. The dissemination westwards of the Indian zero as an integral part of the Indian numerals is one of the most remarkable episodes in the history of mathematics and the story is well-known.

Keywords: numeration system, place value, zero, India

Introduction

If zero merely signified a magnitude or a direction separator (i.e. separating those above the zero level from those below the zero level), the Egyptian zero, *nfr,* dating back at least four thousand years, amply served these purposes. If zero was merely a place-holder symbol indicating the absence of a magnitude at a specified place position (such as, for example, the zero in 101 indicates the absence of any "tens" in one hundred and one), then such a zero was already present in the Babylonian numeration system long before the first recorded occurrence of the Indian zero (Joseph, 98-99). If zero was represented by just an empty space within a well-defined positional numeration system, such a zero was present in Chinese mathematics a few centuries before the Indian zero. The absence of a symbol for zero in China did not prevent it from being properly integrated into an efficient computational tool that could even handle solution of higher degree order equations involving fractions (Joseph, 145) However, the Indian zero alluded to in the question was a multi-faceted mathematical object: a symbol, a number, a magnitude, a direction separator and a place-holder, all in one operating within a fully established positional numeration system. Such a zero occurred only twice in history –the Indian zero which is now the universal zero and the Mayan zero which occurred in solitary isolation in Central America at the beginning of the Common Era.[1]

To understand the first appearances of the Indian and Mayan zeroes, it is necessary to examine them both within the social contexts in which these independent inventions occurred. At the same time we should attempt to identify certain common threads in both cultures that led to the occurrence of the zero in only these two cultures. The dissemination of the Indian zero as a part and parcel of the Indian numerals is one of the most remarkable episodes in the history of mathematics. But what is rarely recognised is that this transmission occurred through a number of cultural and linguistic filters that may have inhibited a clearer understanding of the concept of zero and the arithmetic of the operations with zero. Because of the popular difficulties with the zero, there has occurred over time a series of avoidance mechanisms to cope with the presence of zero which have far-reaching pedagogical implications. And these include the general absence of any discussion at the educational level of the topic of 'calculating with zero' (*'shunya ganita'*) which was emphasized in practically all Indian texts on mathematics from the time of

1. It is important in this context to recognise the fact that a place value system can exist without the presence of a symbol for zero. The Babylonian and the Chinese numeration systems were good examples. But the zero symbol as part of a system of numerals could never have come into being without a place value system. In neither the Egyptian nor Greek nor the Aztec cultures was there a place value system. A zero as a number in any of these systems would in any case have been superfluous (Menninger, 391-392).

Brahmagupta (b. 598 AD) onwards. This is a serious deficiency in the mathematics curriculum both in schools and colleges and needs urgent rectification. As illustrations of this deficiency, consider how the following questions will be answered by students of mathematics:

1. Is zero a positive or negative number?
2. Is zero an odd or even number?
3. Divide 2 by zero

It is not uncommon to find that even among the university students of mathematics, a discussion of these three questions tend to be confused. And it is my experience that an approach through history provides an effective and interesting way of introducing this difficult subject.

The History of Zero: The Indian Dimension

The word 'zero' comes from the Arabic '*al-ṣifr'*. *Ṣifr* in turn is a transliteration of the Sanskrit word "*shunya*" meaning void or empty which became later the term for zero. Introduced into Europe during Italian Renaissance in the 12th century by Leonardo Fibonacci (and by Nemorarius a less known mathematician) as '*cifra'* from which emerged the present 'cipher'. In French, it became '*chiffre'*, and in German *'ziffer'*, both of which mean zero.

The ancient Egyptians never used a zero symbol in writing their numerals. Instead they had a zero to represent a value or magnitude. A bookkeeper's record from the 13th Dynasty (about 1700 BC) shows a monthly balance sheet for items received and disbursed by the royal court during its travels. On subtracting total disbursements from total income, a zero remainder was left in several columns. This zero remainder was represented by the hieroglyph, *nfr*, which also means beautiful, or complete in ancient Egyptian. The same *nfr* symbol also labeled a zero reference point for a system of integers used on construction guidelines at Egyptian tombs and pyramids. These massive stone structures required deep foundations and careful leveling of the courses of stone. A vertical number line labeled the horizontal leveling lines that guided construction at different levels. One of these horizontal lines, often at pavement level, was used as a reference and was labeled *nfr* or zero. Horizontal leveling lines were spaced 1 cubit apart. Those above the zero level were labeled as 1 cubit above *nfr*, 2 cubits above *nfr* and so on. Those below the zero level were labeled 1 cubit, 2 cubits, 3 cubits, and so forth, below *nfr*. Here zero was used as a reference for directed or signed numbers.

It is quite extraordinary that the Mesopotamian culture, more or less contemporaneous to the Egyptian culture and who had developed a full positional value numeration system on base 60 did not use zero as a number. A symbol for zero as a place-holder appeared late in the Mesopotamian culture. The early Greeks, who were the intellectual inheritors of Egyptian mathematics and science emphasised geometry to the exclusion of everything else. They did not seem interested in perfecting their number notation system. They simply had no use for zero. In any case, they were not greatly interested in "arithmetic, claiming that arithmetic should only be taught in democracies for it dealt with relations of equality". On the other hand, geometry was the natural study for oligarchies for "it demonstrated the proportions within inequality. "

In India, the zero as a concept probably predated zero as a number by hundreds of years. The Sanskrit word for zero, *shunya*, meant "void or empty". The word is probably derived from *shuna* which is the past participle of ***svi*,** "to grow". In one of the early Vedas, *Rigveda,* occurs another meaning: the sense of "lack or deficiency". It is possible that the two different words, were fused to give '*shunya*' a single sense of "absence or emptiness" with the potential for growth. Hence, its derivative, *Shunyata*, described the Buddhist doctrine of Emptiness, being the spiritual practice of emptying the mind of all impressions. This was a course of action prescribed in a wide range of creative endeavours. For example, the practice of *Shunyata* is recommended in writing poetry, composing a piece of music, in producing a painting or any activity that come out of the mind of the artist. An architect was advised in the traditional manuals of architecture (the *Silpas*) that designing a building involved the organisation of empty space, for "it is not the walls which make a building but the empty spaces created by the walls." The whole process of creation is vividly described in the following verse from a Tantric Buddhist text:

> "First the realisation of the void (*shunya*),
> Second the seed in which all is concentrated
> Third the physical manifestation
> Fourth one should implant the syllable"

The mathematical correspondence was soon established. "Just as emptiness of space is a necessary condition for the appearance of any object, the number zero being no number at all is the condition for the existence of all numbers" (Datta, 1927).

A discussion of the mathematics of the *shunya* involves three related issues: (i) the concept of the *shunya* within a place-value system, (ii) the symbols used for *shunya*, and (iii) the mathematical operations with the *shunya*. Material from appropriate early texts are used as illustrations below.

It was soon recognised that the *shunya* denoted notational place (place holder) as well as the "void" or absence of numerical value in a particular notational place. Consequently all numerical quantities, however great they may be could be represented with just ten symbols. A twelfth century text (*Manasollasa*) states:

> "Basically, there are only nine digits, starting from 'one' and going 'nine'. By the adding zeros these are raised successively to tens, hundreds and beyond."

And in a commentary on Patanjali's *Yogasutra* there appears in the seventh century the following analogy:

> "Just as the same sign is called a hundred in the "hundreds" place, ten in the "tens" place and one in the "units" place, so is one and the same woman referred to (differently) as mother, daughter or sister."

The earliest mention of a symbol for zero occurs in the *Chandahsutra* of Pingala (fl. 3rd century BC) which discusses a method for calculating the number of arrangements of long and short syllables in a metre containing a certain number of syllables (ie. the number of combinations of two items from a total of *n* items, repetitions being allowed). The symbol for *shunya* began as a dot (*bindu*), found in inscriptions both in India and in Cambodia and Sumatra around the seventh and eighth century and then became a circle (*chidra* or *randra* meaning a hole). The association between the concept of zero and its symbol was already well-established by the early centuries of the Christian era, as the following quotation shows:

> "The stars shone forth, like zero dots (*shunya-bindu*) - scattered in the sky as if on the blue rug, the Creator reckoned the total with a bit of the moon for chalk." (Vasavadatta , ca 400 AD)

Sanskrit texts on mathematics/astronomy from the time of Brahmagupta usually contain a section called '*shunya-ganita'* or computations involving zero. While the discussion in the arithmetical texts (*patiganita*) is limited only to the addition, subtraction and multiplication with zero, the treatment in algebra texts (*bijaganita*) covered such questions as the effect of zero on the positive and negative signs, division by zero and more particularly the relation between zero and infinity (*ananta*) (Pandit, 1990).

Take as an example, Brahmagupta's seventh century text *Brahmasphuta-Siddhanta*. In it, he treats the zero as a separate entity from the positive (*dhana*) and negative (*rhna*) quantities, implying that *shunya* is neither positive nor negative but forms the boundary line between the two kinds, being the sum of two equal but opposite quantities. He stated that a number, whether positive or negative, remained unchanged when zero is added to or subtracted from it. In multiplication with zero, the product is zero. A zero divided by zero or by some number becomes zero. Likewise the square and square root of zero is zero. But when a number is divided by zero, the answer is an undefined quantity "that which has that zero as the denominator." (Datta, 1927).

The earliest inscription in India of a recognisable antecedent of our numeral system is found in an inscription from Gwalior dated '*Samvat* 933' (876 AD).[1] The spread of these numerals westwards is a fascinating story. The Islamic world was the leading actor in this drama. Indian numerals probably arrived at Baghdad in 773 AD with

1. There is earlier evidence of the use of Indian system of numeration in South East Asia in areas covered by present-day countries such as Malaysia, Cambodia and Indonesia, all of whom were under the cultural influence of India. Also, as early as 662 AD, a Syrian bishop, Severus Sebokt, comments on the Indians carrying out computations by means of nine signs by methods which "surpass description" (Joseph, 311-312).

the diplomatic mission from Sind to the court of Caliph al-Mansur. In about 820 al-Khwārizmī wrote his famous *Arithmetic*, the first known Arabic text to deal with the new numerals. The text contains a detailed exposition of both the representation of numbers and operations using Indian numerals. Al-Khwārizmī was at pains to point out the usefulness of a place-value system incorporating zero, particularly for writing large numbers. Texts on Indian reckoning continued to be written and by the end of the eleventh century, this method of representation and computation was widespread from the borders of Central Asia to the southern reaches of the Islamic world in North Africa and Egypt.

In the transmission of Indian numerals to Europe, as with almost all knowledge from the Islamic world, Spain and (to a lesser extent) Sicily played the role of intermediaries, being the areas in Europe which had been under Muslim rule for many years. Documents from Spain and coins from Sicily show the spread and the slow evolution of the numerals, with a landmark for its spread being its appearance in an influential mathematical text of medieval Europe, *Liber Abaci*, written by Fibonacci (1170-1250) who learnt to work with Indian numerals during his extensive travels in North Africa, Egypt, Syria and Sicily.[1] And the spread westwards continued slowly, displacing Roman numerals, and eventually, once the contest between the abacists (those in favour of the use of abacus or some mechanical device for calculation) and the algorists (those who favoured the use of the new numerals) had been won by the latter, it was only a matter of time before the final triumph of the new numerals occurred with bankers, traders and merchants adopting the system for their daily calculations (Joseph, 312-316).

1. There is a tendency to concentrate on the contribution of Fibonacci in the spread of the Indo-Arabic numerals into Europe. But there were other disseminators as well. When it came to Scandinavia the book of Hauk was of critical importance. Entitled *Algorismus*, it began: “This art … was first discovered by the Indians (who) used ten figures written like this 0 9 8 7 6 5 4 3 2 1. The first number is one, the second two, the third three and so forth, until the last which is called cifra. And these symbols begin from right and are written to the left in the manner of the Hebrews…. Cifra doesn’t count on its own but gives place and hence meaning to other figures.”

The History of Zero: The Mayan Dimension

Evidence relating to Pre-Columbian Maya civilisation comes from three main sources: four screen-fold books called codices, a large number of stone monuments and thousands of ceramic vessels. The best account of the Maya culture around the time of the Spanish Conquest comes from a Franciscan priest, Diego de Landa who recorded the history and traditions of the Maya people around 1566. Piecing together these different strands of evidence, it is possible to construct an account of the social context in which the Mayan numerals and especially the Mayan zero emerged around the beginning of the Christian era.

The Mayan sytem of numerical notation was one of the most economical systems ever devised. In the form that was used mainly by the priests for calendar computation as early as 400 BC, it required only three symbols: a dot was used for one, and a bar for five; and a symbol for zero which resembles a snail's shell. With these three symbols they were able to represent any number on a base 20. However, there was an unusual irregularity in the operation of the place value system. Corresponding to our units, tens, hundreds, thousands, ….. etc, the Mayans had units, 20's, (18×20)'s, (18×20^2)'s, (18×20^3)'s, …. etc. This anomaly reduces the efficiency in arithmetical calculation. For example, one of the most useful facilities in our numeration system is the ability to multiply a given number by 10 by adding a zero to the end of it. An addition of a Mayan zero to the end of a number would not in general multiply the number by twenty because of the mixed base system employed. This inconsistency also inhibited the development of further arithmetical operations, particularly those involving fractions.

To understand this curious irregularity in Mayan numeration, it is important to appreciate the social context in which the numeration system was used. As far as we know this form of writing numbers was used only by a tiny elite – a group of priest scribes who were responsible for carrying out astronomical calculations and constructing calendars. At the top of the pyramid was a hereditary leader who was both a high-priest (*Ahau-Can*) and a Maya noble. Under him were the master scribes who were priests as well as teachers and writers (engaged in teaching their sciences as well as in writing books about them). Mathematics was recognised as such an important discipline that depictions of scribes who were adept at that discipline appear in the iconography of Mayan artists. Their mathematical identity was signified in the manner in which they were depicted: either with the Maya bar and dot numerals coming out of their mouths or a number scroll being carried under their armpit. The location of the scroll under the armpit with numbers written on it would seem a status symbol. In an interesting illustration on another Maya vase from the beginning of the Christian era, there is a seated supernatural figure with the ears and hooves of a deer, attended by a number of human figures, including a kneeling scribe mathematician from whose armpit emanates a scroll containing the sequence of numbers 13, 1, 3, 3, 4, 5, 6, 7, 8 and 9. At the top right hand corner of this illustration there is the small figure of a scribe who looks female, with a number scroll under her armpit indicating that she is a mathematician and possibly the one who painted the scene and wrote the text on the vase. She is described as *Ah T'sib* (the scribe). Preceding this text is a glyph that has not been deciphered but which could be her name. Once the name is deciphered, and if the scribe is female, we may have the name of one of the earliest known women mathematician-scribe in the world. The existence of female mathematician/scribes among the Maya is further supported by another depiction found on another ceramic vase. The text on this vessel contains the statement of the parentage of the scribe in question: "Lady Scribe Sky, Lady Jaguar Lord, the Scribe". Not only does she carry the scribal title at the end of her name phrase but she incorporates it into one of her proper names, an indication of the important role she herself plays on that reality (Joseph, 367-368).

Returning to the curious irregularity in the Mayan place value system, the general view is that it is tied to the exigencies of operating three different calendars. The first calendar, known as the *tzolokin* or ‘sacred calendar’, contained 260 days in twenty cycles of 13 days each. Superimposed on each of the cycles was an unchanging series of twenty days , each of which was considered a god to whom prayers and supplications were to be made. The second, known as a civil or secular calendar, was the one for practical use. It was a solar calendar consisting of 360 days grouped into 18 monthly periods of twenty days and an extra month consisting of five days. The last month was shown by a hieroglyph that represented disorder, chaos and corruption and any one born in that month was supposed to have been cursed for life. Finally, there was the third calendar of 'long counts' similar to the Indian *'Yuga'* periodisation. The upper section of one of the oldest standing stelas at Ires Zapotes in Mexico shows the date of its construction in the calendar of 'longcounts' as:

8 kins = 8 × 1 =	8 days	(20 kins = 1 uinal)
16 uinals = 20 × 16 =	3 20 days	(18 uinal = 1 tun)
0 tuns = 20 × 18 × 0 =	0 days	(20 tuns = 1 katun)
6 katuns = $(18)20^2 \times 6$ =	43/200 days	(20 katuns = 1 baktun)
16 baktuns = $(18)20^3 \times 16$ =	304000 days	(20 baktuns = 1 piktun)
7 piktuns = $(18)20^4 \times 7$ =	20160000 days	(20 piktuns = 1 calabtun)
TOTAL	22,507,528 **days** which corresponds to **31 BC**[1]	

There were higher units of measurement, notably ***kinchiltuns*** and ***alautins*** where 1 ***alautin*** equalled 23,040,000,000 days. Measurement of time constituted a central feature of the Mayan culture and the interest in measurement was carried into Mayan astronomy. We can only marvel at the high degree of accuracy that the Mayans achieved in their astronomical work. To illustrate, without any sophisticated equipment and with the deficiency of a mixed base system, they obtained the mean duration of a solar year as 365.242 days (modern value: 365.242198 days) and the mean duration of a lunar month as equivalent to 29.5302 (modern value: 29.53059 days).

The Two Zeroes: Common Threads and Differences

I began this paper with the question relating to the Indian zero which has now been extended to include the Mayan Zero. Why did the

---

1. The start date of the Mayan 'long count calendar', expressed in terms of our calendar, was Augest 13 in the year 3114 BC. Their calendar had an end date of December 23 of 2012 AD when there was supposedly some enormous catastrophe that might even mark the end of the world!

full use of zero within a well-established positional value system only emerge in two cultures. Were there any similarities between the two cultures that might provide an answer, however tentative it remains.

From the existing evidence, much of it fairly fragmentary especially in the Mayan case, we are aware that both cultures were numerate with considerable interest in astronomy. The Indian culture from an early time showed interest and even fascination for large numbers and there is no contrary evidence to indicate that this was not so in the Mayan culture. Both cultures were obsessed with the passage of time but in different ways. The Indian interest was tied up the wide-spread belief in a never-ending cycle of births and rebirths with the primary objective for individual salvation being the need to break the cycle. This was apparently achieved during the Vedic times by carrying out sacrifices on specially constructed altars which conformed to specific shapes and sizes and where the sacrifices had to be carried out on particular days chosen for their astronomical significance. In the Mayan case, the obsession took the form of a society's fear that the world would come to an end unless the gods (and especially the Sun God) were propitiated by human sacrifice to be undertaken systematically at certain propitious time of the year to be dictated by specific astronomical occurrences. In both cases there was need for accurate measurement of time and hence the detailed calendars and the elaborate periodisation into eras. The need for such precise calculations may have stimulated the development of efficient numeration systems with a fully developed zero. And it was probably only an accident of history and geography that the Indian zero prevailed while the Mayan zero eventually disappeared into oblivion (Teresi, 80-84).

References

Arsham, H., "A Critical Panoramic View of Basic Mathematical Concepts", *Teaching Mathematics and Its Applications* 17, No. 2, 1998, pp.69-72.

Datta, B., "Early history of the Arithmetic of Zero and Infinity in India", *Bulletin of the Calcutta Mathematical Society* 18, 1927, pp.165-176.

Joseph, G.G., *The Crest of the Peacock: Non-European Roots of Mathematics*, Princeton University Press, Princeton and London, 2000.

Yong, Lam Lay, "Linkages: Exploring the similarities between the Chinese rod numeral system and our numeral system", *Archives for History of Exact Sciences* 37, 1987, pp.365-392.

Lumpkin, B., "Mathematics used in Egyptian Contstruction and Bookkeeping", *The Mathematical Intelligencer* 24, no. 2, 2002, pp. 20-25

Menninger, K., *Number Words and Number Symbols*, Dover Publications, New York, 1992.

Pandit, M.D., *Zero in Panini*, CASS Publications, University of Pune, India, 1990.

Teresi, D., *Lost Discoveries*, Simon & Schuster, New York, 2002.