

A Quick Introduction to Indian Manuscripts for the Non-Specialist

The rich array of manuscript material from South Asia and its surroundings is something little studied in the West, outside of very specialized researchers, and not without reason. The bewildering array of languages, scripts, and formats which one encounters can be off-putting at best, and anyone having to sort and process a collection of this material faces a daunting task. Even specialists can be swamped, as so much of the material is outside any possible specialty, so that an expert on, say, Vedic commentaries would hardly be likely to have much expertise in Persian epic poetry or Theravada Buddhist subcommentaries on the Abhidhamma¹, for example. As a paper researcher who has had to collect his own specimens I have encountered this situation, so I am familiar with the problems; and as a paper researcher with only the most superficial skills with the languages involved – I catalog coins, so I have enough to puzzle out inscriptions – I cannot read very much, but I still have to know enough about the documents to be able to guess about where and when they were written. This paper is an attempt at summarizing what I have found out about manuscript material from India – and some of the surrounding areas – from thirty-odd years of working with them. Hopefully it will help to enable the generalist to begin to be able to sort out a mass of this material, and get it into the hands of the specialists who need to see it.

1. Formats

The first thing anyone notices when looking at a miscellaneous group of these MSS is the variety of formats. However complex they appear at a glance, however, there are two main formats, derived from

¹ To pull a subject out of the air; like the joke about the fellow who warns off attackers, because he “knows karate, kung-fu, tae-kwan-do, aikido, and five other Chinese, Japanese, and Korean words”.

two main traditions: first, the native palm-leaf format; and, second, the imported Western-style codex format. Everything else is a derivative of one of these two. Occasionally we may encounter something a bit different, such as the folding books from Thailand and Burma, for example; but these are still derived ultimately from palm-leaves.

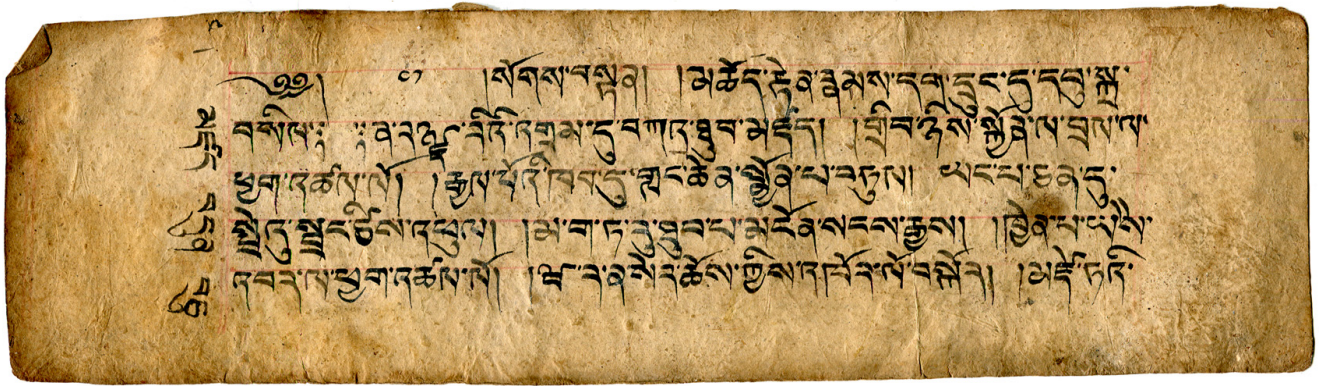


Illustrated palm-leaf from a fragmentary copy of the Gitagovinda, Orissa, XX century.

Throughout India various sorts of palm-leaves provided the main writing material from earliest times. In northern India the palm-leaf was replaced by paper², at different times in different areas, between the fourteenth century (Gujarat and vicinity) and the sixteenth (Nepal); in southern India and surrounding areas (Thailand, Burma, Lanka) palm-leaves remained the primary writing material into the twentieth century. On the other hand, in Tibet, paper became the primary writing material more or less with the introduction of Buddhism in the seventh century, and the only major change was the general replacement of the native paper with imported paper in the course of the nineteenth century³.

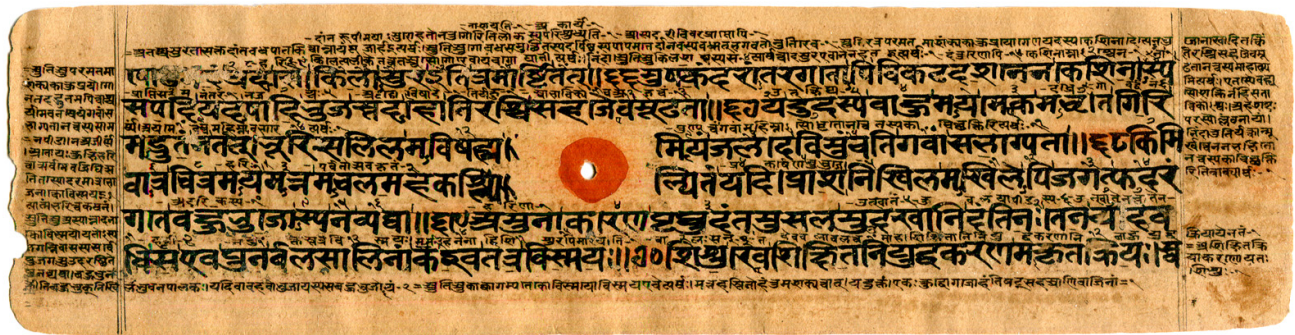
2 And in some cases birch-bark; such MSS exist, but they are rarely met with. They are usually bound in the single-quire format; see below.

3 Mongolia as well, as the Mongols converted to Tibetan Buddhism in the sixteenth century. The sequence of paper is of interest, especially as the material is (at the moment) readily available in some quantity.



Tibetan MS leaf, unidentified text, XVIII-XIX century

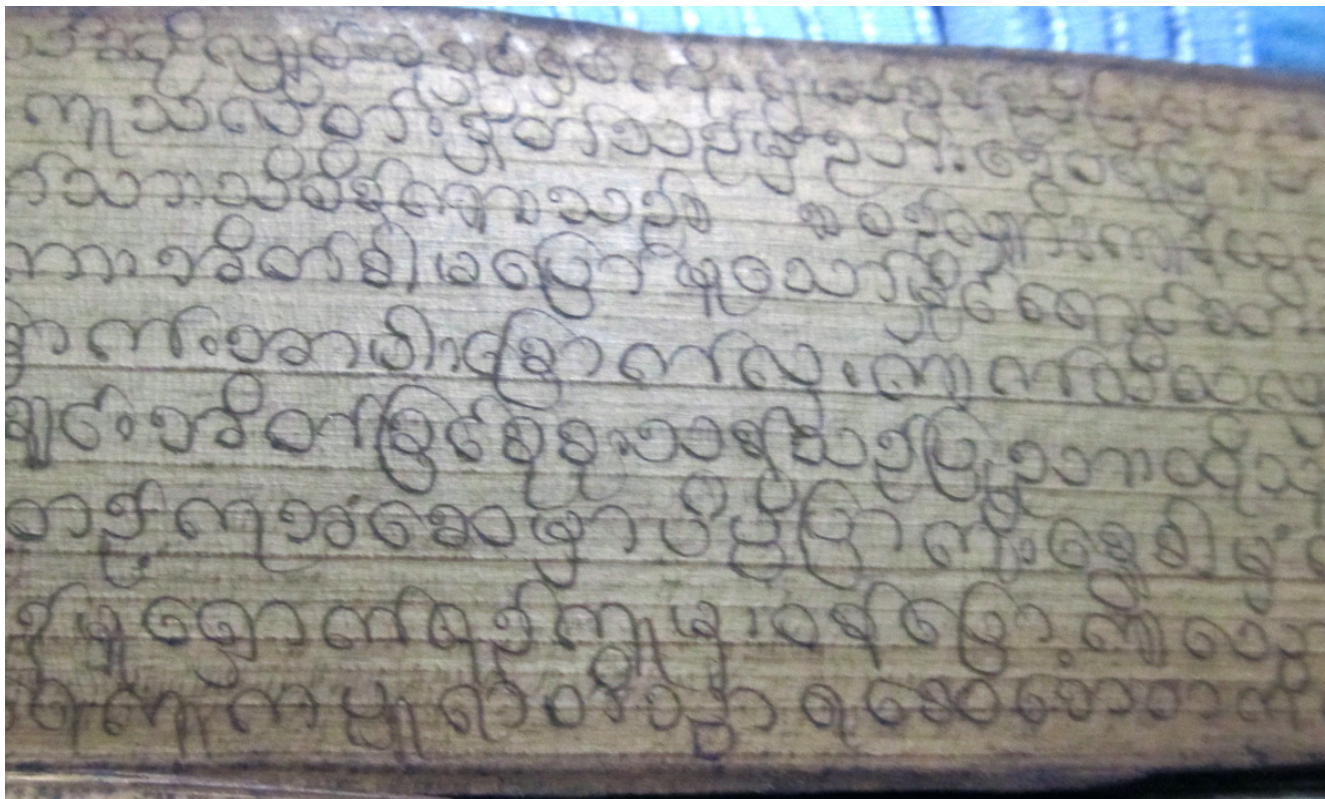
The paper which replaced the palm-leaves generally followed the same format as the leaves, namely long and narrow, with text which read properly when the leaf was turned vertically. This was the most common format for Hindu and Jain MSS through the nineteenth century. Nonetheless there were many others, some simply of different dimensions, while others were adapted slightly to the codex format. Some of these were simply loose leaves which turned horizontally rather than vertically; others were bound along the top, or along the left edge, like long skinny codices.



Uncertain Jain text on paper, n palm-leaf format, XIV-XV Century.

In northern India the leaves were written on using pen and ink. In southern India the writing was inscribed into the leaves with a stylus, and ink later rubbed in. This accounts for the differences in

the style of scripts; the long horizontal line (or series of short horizontal lines) at the top of the aksharas⁴ is not a problem if written in ink, but if inscribed on a leaf will often split it. Therefore wherever a stylus is used the characters are rounded; the ultimate is the Burmese script, which can look like a random series of circles. All of the scripts are ultimately derived from the same Brahmi script, however different they may look.



Example of the ultimate round script, Burmese.

The other format is the Western-style codex, introduced in its present form by Muslim conquerors, most probably the Ghorids who established the Sultanate of Delhi in 1192. It was the usual

4 The term akshara is used rather than “character” or “letter” because the Brahmi script and its derivatives are syllabic, and occasionally compound, rather than letters as we are used to seeing. The aksharas are usually a consonant plus a vowel (e. g. sa); sometimes a consonant plus a vowel plus a consonant (sam); sometimes two or more consonants plus a vowel (stha); sometimes consonant(s) without a vowel (sth); and sometimes a consonant plus two blended vowels (ta + a = t with long a).

format for books in Arabic, Persian, and, later, Urdu languages. Much later it was adopted by the Sikhs, who with their synthesis of Islam and Hinduism used Islamic book formats. Later still several variants were adopted by Hindus and Jains, but with major differences. A Persian binding, for example, is made up of a number of discrete quires; a typical Hindu or Jain codex is usually a single, thick quire, or else a group of loose leaves which turn horizontally, like the leaves of a normal codex; these are occasionally sewn together for convenience of use. Exceptions do occur, of course; the Hindu manuscripts of Kashmir with their ornate Sarada script and decorated borders are usually made up of many discrete quires as well. The main difference between these and the Persian MSS lies in the binding; Persian MSS are sewn onto cords through the folded back of the quire, like Western MSS; Kashmiri MSS, on the other hand, are stitched through the inner margins of the leaves, as befits a tradition in which the original MSS consisted of loose leaves stitched together along the left margin.

2. Paper

As my specialty is paper, I must of course make a few remarks on it. The subject has been little studied, so any comments are at best speculative; but with continued investigation things are beginning to fall into place. It is now possible to begin to work up a typology of paper, showing changes over both space and time. Here we can only touch on a few of the high points.

Watermarked paper in the Subcontinent is unusual and nearly always quite late. Occasionally we get a sample with the usual Italian marks; such a specimen was most likely brought in from elsewhere. Other than that we find obsolete English legal-document paper, beginning in the 1860s, in some cases. Although this paper usually is dated, it was by definition imported only after it had become obsolete for its original function, so the date of the document is invariably later than the date of the watermark. This

is also true enough for the Russian paper found in Mongolia, where it becomes the usual paper found after ca. 1790.

The local product, on the other hand, is never found with watermarks, and rarely even has chain lines; when it does have chain lines, they are inconsistent and randomly spaced, with a few exceptions. The only really useful criterion found in the wire-lines⁵ is that the laid lines vary, from very coarse and straight in some of the earlier examples to very fine and straight in some later ones to fairly coarse and curved in yet later examples. But dated examples of early material are very scarce, and most of the dated material we have is Moghul or later.

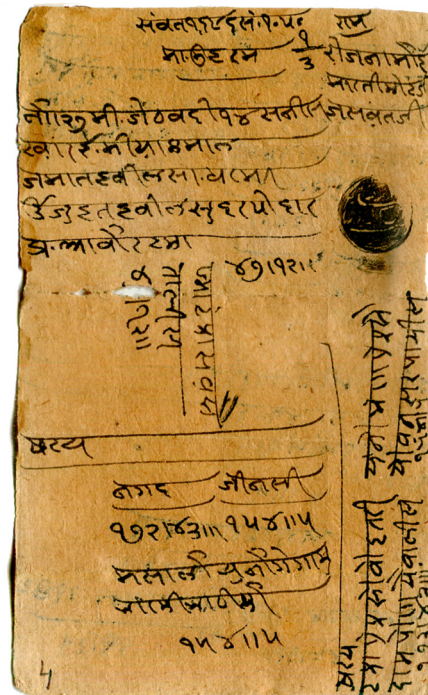


Example of wire-lines in a Jain MS, XVI Century.

Long sequences of dated material are uncommon; but we do have one of note. The state of Jaipur, originally called Amber from the original capital city, produced a large volume of tax documents

⁵ The conventional term; as has been pointed out by e. g. Loveday the moulds used in India use grasses or reeds rather than wire for the screens.

which were scattered after the state was absorbed by the Republic in 1949, and are frequently available very cheaply from stamp dealers. I have a considerable number of them, and those I have seen range from the 1630s to the 1890s. They therefore provide an excellent, if geographically limited, paper sequence. The main difference between earlier and later paper is the colour; the appearance of the laid lines evolves as well.



Example of an Amber/Jaipur tax document, here dated Samvat 1696, AH 1050 (AD 1639/40).

3. Languages and scripts

On the Subcontinent we find a bewildering array of scripts and languages which can look very daunting to the neophyte. They are in many ways not as bad as they look. There are two basic scripts within India, including Southeast Asia and Tibet. The first is the Arabic script; the second is the Brahmi script, each with its derivatives. Generally speaking the first is Islamic, and the second covers everything

else. To work with any of this material we need decent reference material, particularly a guide to scripts. For beginners the best choice is probably the ALA-LC Romanization Tables, available from Amazon as “print-on-demand” as it seems to be out of print⁶, and is no longer available from the ALA website⁷. Online resources can be helpful; most of them are superficial (e. g. AncientScripts.com)⁸, or very dense (indoskript.org)⁹. The problem with most of these resources is that they usually only cover the modern version; the scripts changed with time, and some of the aksharas encountered in manuscripts don’t match anything very well, hence the use of Indoskript.

Regarding Arabic script, there are two main forms, namely regular and cursive. The former includes a lot of calligraphic variants, which all have in common that the letters are all spelled out, whether the script is Naskh, Bihai, Thuluth, monumental Kufic, or some other variety; the cursive type, in which the letters are often abbreviated, is usually some form of Nastaliq¹⁰. The latter is usually used for Persian or Urdu; the former, for Arabic; but Naskh was often used, particularly in earlier MSS, for Persian as well. Most of the calligraphic scripts in MSS were used for writing Qurans, and occasionally for other works. Occasionally a regional language appears in Arabic script, with its own additional letters, which is one way of identifying the various languages¹¹.

6 https://www.amazon.com/ALA-LC-romanization-tables-Transliteration-non-Roman/dp/B002VWLKQ8/ref=tmm_pap_title_0?encoding=UTF8&qid=&sr= (3 March 20)

7 https://www.alastore.ala.org/search/store/subject/cataloging-knowledge-and-information-management?filter_all=1&items_per_page=72 (3 March 20)

8 <http://www.ancientscripts.com/index.html> (3 March 20)

9 <http://www.indoskript.org/> (3 March 20)

10 An excellent summary of script varieties is found in Blair, Sheila; *Islamic Calligraphy*, Cairo 2006.

11 e. g., Persian has letters for P and Ch, and Urdu has letter equivalents for all the Sanskrit letters; the occasional Turkish, Afghani, Punjabi, or Kashmiri MS will be a similar case.

مثله للشعبي فزاده حُسْنًا أَخْبَرَنَا عُمَرُ بْنُ أَحْمَدَ الْوَاعِظُ
 حَدَّثَنَا مُحَمَّدُ بْنُ الْحُسَيْنِ حَدَّثَنَا مُحَمَّدُ بْنُ الْحَارِثِ حَدَّثَنَا جَدِّي
 حَدَّثَنَا الْهَيْثَمُ بْنُ عَبْدِ حَيْدٍ حَدَّثَنَا مُحَمَّدُ بْنُ خَالِدٍ عَنِ الشَّعْبِيِّ قَالَ تَعَاشَرُوا
 النَّاسَ زَمَانًا طَوِيلًا بِالذِّينِ حَتَّى ذَهَبَ الدِّينُ • ثُمَّ تَعَاشَرُوا
 بِالْمُرُوءَةِ حَتَّى ذَهَبَتِ الْمُرُوءَةُ • ثُمَّ تَعَاشَرُوا بِالْحَيَاءِ حَتَّى
 ذَهَبَ الْحَيَاءُ • ثُمَّ تَعَاشَرُوا بِالرَّغْبَةِ وَالرَّهْبَةِ وَأَطْنَبَهُ
 سَيِّئَاتِي بَعْدَ ذَلِكَ مَا هُوَ شَرُّ مِنْهُ • وَمِنْ آدَابِ الصَّحْبَةِ

Example of ordinary Naskh script.

که گوهر اصلا به کند و به او دهند و پادشاه میفرماید که گوهر را زود بشکن
 و صلابه کرده پیاور • تمام شد این ارقام پریشان پهن تمة درویشان
 در تاریخی رفیع الشان فیض نشان صغر تسعین و شانمانه از هجره خاتم
 صلی الله علیه و سلم موافق اسفند پسته و اربعانه جلالی که زمان
 قرآن بر چسب و کیوانست در عقرب و او ان اجتماع سبعة سیاره آن
 غیر بهرام در برج سعاده انجام امید آنکه از غیر محرم محفوظ و از نظر
 مخطوط باشد در آخرة واسطه خلاص و نجات و رابطه رفعة در جات

Example of Persian Nastaliq script

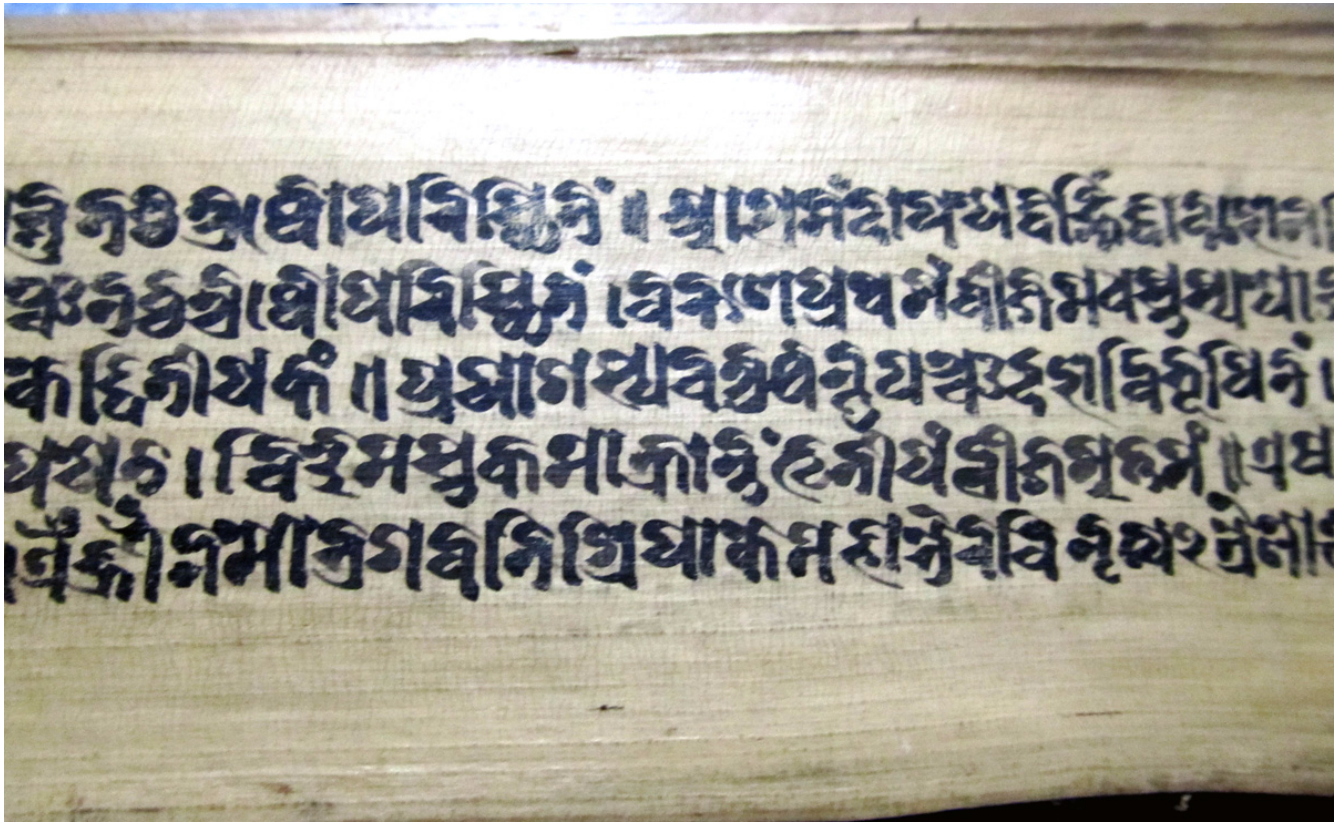
Brahmi scripts of various sorts were used in India since earliest recorded times. We, however, are very unlikely to run into anything really early; the earliest MSS which turn up are fragments of Bamiyan palm-leaf MSS imported before the import ban, mostly by a book and MS dealer from Akron, OH. They are Buddhist in origin, and mostly date to around 500 AD, and the language is Sanskrit and the script one of the Gupta varieties of Brahmi.



Unidentified Buddhist text fragment on palm-leaf, northern India, c. VI century.

The next earliest sort of material is from Nepal, on palm-leaves, mostly between 1100 and 1500 AD; it is also usually Buddhist, but occasionally Hindu. The language is Sanskrit, and the script is now

Devanagari or a variant (i. e., either a calligraphic variant such as Bhujjmol or a regional variant such as Newari). The script differences are either decorative, or regional.



Example of Bhujjmol style of Nepali script; Nepal late XII Century.

Space prohibits more than a quick glance at regional differences. In the west we have Kashmiri, written with the “Sarada” script, which branched off from Nagari in c. 9th century.



Example of formal Sarada script. Kashmir. XIX century.

Punjabi, also called Gurumukhi, used mainly by the Sikhs, was a deliberate construction beginning with the religion itself in the sixteenth century. Gujarati, nowadays distinguished by the lack of a top line for the aksharas, lost the line only recently. The Jains in western India wrote much of their literature in Jain Prakrit, which is close to Sanskrit; they used Devanagari for some literature, and a local variety called “Jain Nagari”¹² for other types.

Most of central India used fairly standard Devanagari script and Sanskrit or Hindi for the language, although many minor variants exist to confuse us. The Jaipur tax documents noted above are said to have been written in a somewhat arcane language by the scribes who wrote them, for example¹³.

¹² <http://www.jainpedia.org/resources/what-is-a-jain-manuscript/contentpage/4.html> (3 Mar 20)

¹³

In eastern India the main variety of script is Bengali, whose angular aksharas are distinctive. It is derived from Buddhist Devanagari; the beginnings can be traced in Nepal.

The inscribed palm-leaf traditions can be fairly quickly summarized. In India proper the only important Indo European language found is Orya, from Orissa, which is a variety of Devanagari script with rounded tops on the aksharas; the literature is almost entirely standard Sanskrit Hindu texts. Most of the rest of the palmleaf material is one of the Dravidian languages, which are mainly Tamil, Telugu, or Malayalam, written in their characteristic scripts or in a general script known as Grantha, from Granth, a book. The literature is mostly Hindu, but occasionally Christian, as Christianity has a long history there, predating the Portuguese arrival by many centuries.

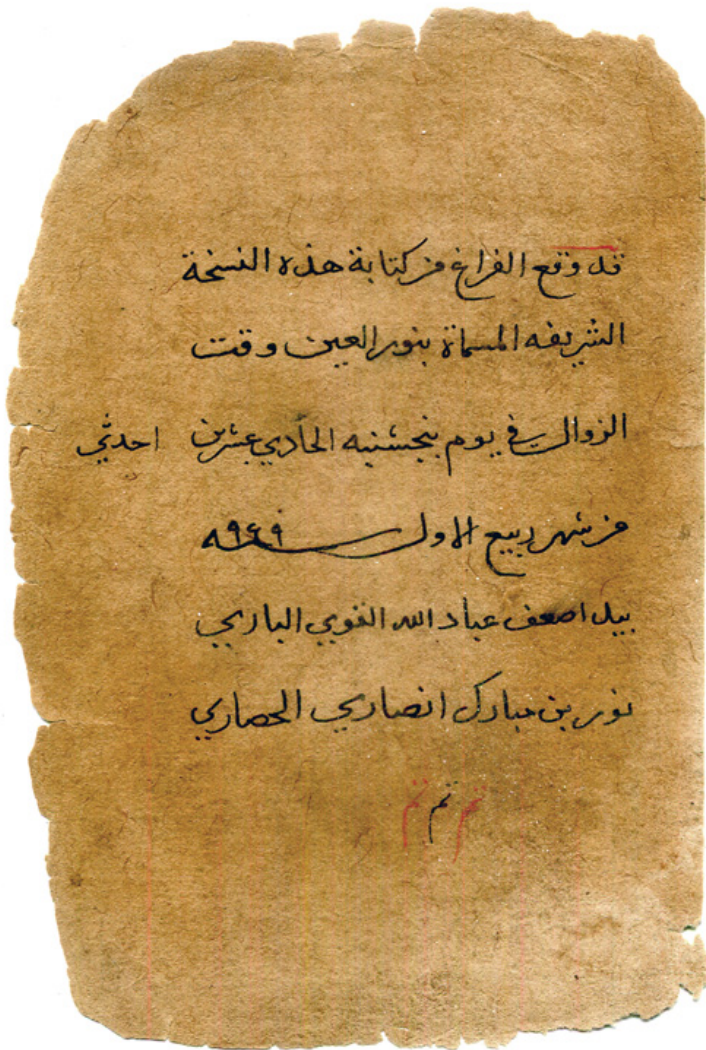
Keeping within the boundaries of the Subcontinent, we only have one major area remaining, the island of Lanka, formerly called Ceylon, a corruption of Sri Lanka. Home to majority Sinhalese, mostly Buddhists, and minority Tamils, mostly Hindus, it has its own format. The Tamil literature is essentially that of the mainland. The Sinhalese is written in either Pali, essentially a Theravada Buddhist dialect of Sanskrit¹⁴, or Sinhala, an Indo-European language not very closely related to anything else in India. It is generally Buddhist, with a lot of secular and semi-secular literature.

4. Colophons and titles

Determining the title of the text, along with its place, date, and copyist, can be a challenge. If the book is intact it can be enough of a challenge. With Arabic and Persian texts it can be a particular challenge, especially when dealing with fragments; the title is typically written on the first leaf, and the rest of the information in the colophon. As Arabic bindings are usually fairly weak they often fall apart

¹⁴ Basically, a sort of “popular” offshoot of Sanskrit used in the early Buddhist territories in the last few centuries BC.

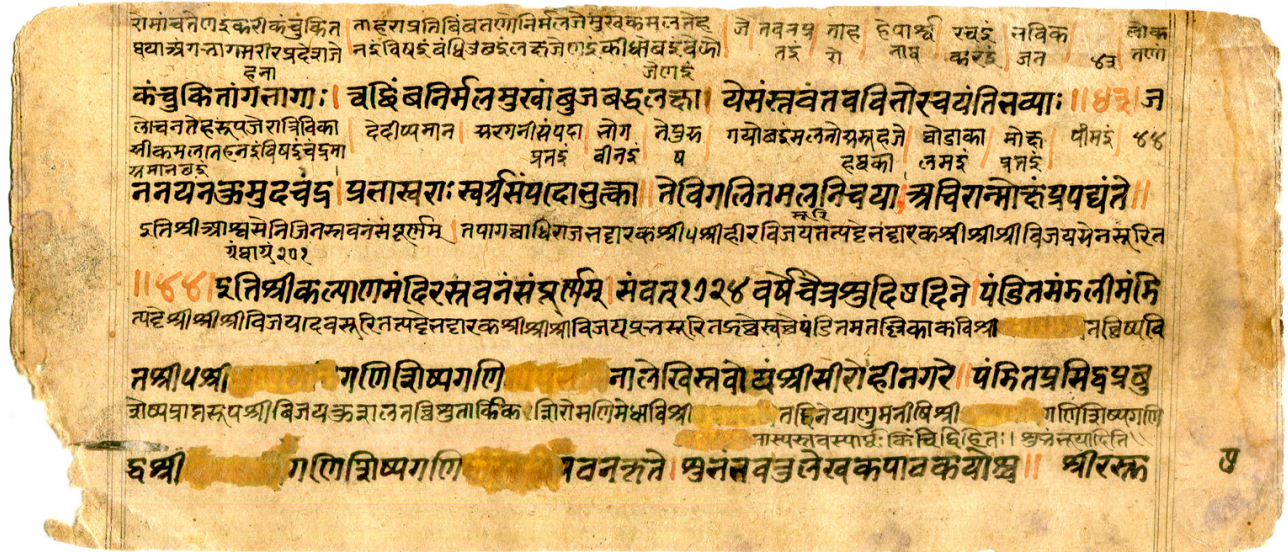
and lose the first few and last few leaves or quires. If one has the colophon it sometimes gives useful information, especially for later MSS. So in the ideal case we might have a colophon reading “This copy of Sahih Bukhari was completed by me ‘Abd al-Rahman ibn Ahmad al-Isfahani in Dar al-Khalifat on 16th Ramadan year four and thirty and two hundred and one thousand (AH, anno hejira)”¹⁵. Many lack some or all of this information. For beginners the date of copying is complicated enough. The general formula is, however, the same whether the text is in Arabic, Persian, or Urdu.



Islamic colophon leaf, giving, among other things, a date of 21 Rab'i al-Awwal AH 959 (AD 1551).

15 The year is usually the usual Islamic era from AD 622 in a lunar calendar of 354 days. Thus year 0 is 621/2; 100=718/9; 200=815/6; 300=912/3; 400=1009/10; 500=1106/7; 600=1203/4; 700=1300/1; 800=1397/8; 900=1494/5; 1000=1592; 1100=1688/9; 1200=1785/6; 1300=1882/3; 1400=1979/80, and so on. Every 30 years AH is 29 years AD, more or less.

Hindu texts in general are much easier if the MS has a colophon, on the last side, and if it gives the required information. Almost all MSS of this sort, whether on palm-leaf or on paper, will end with a title formula: “This is Sri Bhagavad Gita”, often with other explanatory material giving place and date – the latter sometimes in two eras – and usually the month and day. The month is usually that named in the Hindu calendar¹⁶; the year is either Vikrama Samvat (=AD + 57) or Saka (=AD - 78). The abovenoted Jaipur tax documents, like most of the Indian legal papers, are dated mainly in VS, occasionally in AH as well. Buddhist literature, mostly from Lanka, is dated by the Buddhist era, which is AD + 543/4¹⁷, with its specific month names; otherwise the formula is very close to the Hindu one.



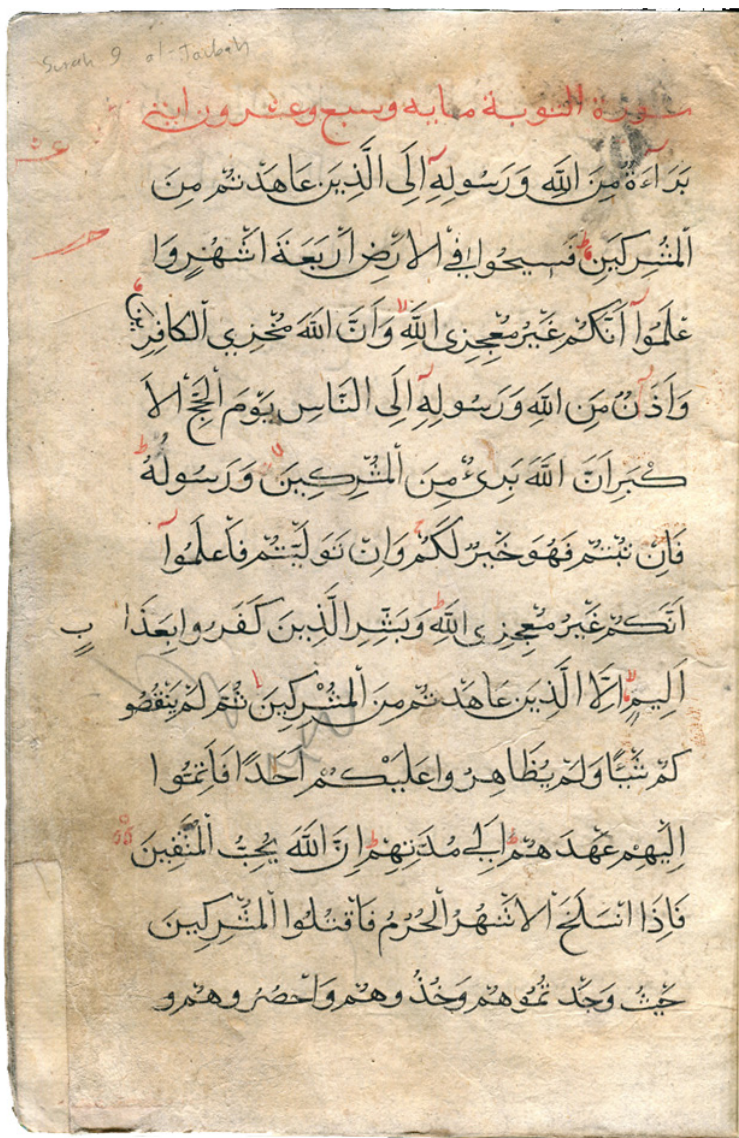
Last leaf of a MS, with colophon on third line of man text containing “Iti sri” and title Kalyanamamdirastavanam followed by date: Samvat 1724 (1668), month Chaitra (Mar/Apr).

5. Contents

16 https://en.wikipedia.org/wiki/Hindu_calendar (3 March 20)

17 https://en.wikipedia.org/wiki/Buddhist_calendar (3 March 20)

Obviously to do more than touch on the literature is far beyond the scope of this paper. Some of it, however, is basic, and can be noted. In Arabic the main text is of course the Quran; with a little practice the cataloger can tell at a glance what it is. Qurans are nearly always accented; they also have chapter headings with the word “surah”, followed by the surah name, then by a number which indicates its place in the chronological order of origin. Nowadays the surahs are arranged from longest to shortest; the names of the surahs can be looked up online¹⁸.



Quran section, Sura 9, Surah at-Taubah. Early Ottoman, XV century.

18 <https://www.sacred-texts.com/isl/htq/> (3 March 20)

الْوَهَّابِ ۝ أَمْرٌ لَهُم مَّلَكُ السَّمَوَاتِ
وَالْأَرْضِ وَمَا بَيْنَهُمَا فَليزْتَقُوا
فِي الْأَسْبَابِ ۝ جُنْدٌ مَّا هُنَالِكَ
مَهْذُومٌ مِنَ الْأَحْزَابِ ۝ كَذَّبَتْ
بِقُلُوبِهِمْ قَوْمٌ نُوْحٌ وَعَادٌ وَفِرْعَوْنٌ
ذُرَّوْا الْأَوْتَارَ ۝ وَثَمُوْدُ وَقَوْمٌ لُوطٌ
وَاصْحَابُ الْأَيْكَةِ ۝ أُولَئِكَ الْأَحْزَابُ ۝
إِن كَلَّ الْأَكْذِبُ الرَّسُلَ
فَخَفَّ عِقَابٌ ۝ وَمَا يَنْظُرُ هُوَ إِلَّا
الْأَيْحَةَ وَاحِدَةً مَّا لَهَا مِنْ
فَوَاقٍ ۝ وَقَالُوا رَبَّنَا عَجَبًا لَنَا
قُطُنًا قَبْلَ يَوْمِ الْحِسَابِ ۝ اجْزِمِ
عَلَيَّ مَا يَقُولُونَ وَإِذْ كُرِعْنَا

داود

Quran in the distinctive Indian Bihari script. XVI century.

The other main texts in Arabic are traditions (hadiths) and Quran commentaries (tafsir). Hadith literature is written to certain formulas (i. e., Aisha related that the Prophet (rasul Allah salih Allah) ...” which we can learn to recognize with a bit of practice. Tafsirs are a bit more complicated; the easiest way is to note an accented Quran quote followed by a section of unaccented text. That can give an idea that it is a Tafsir; but there are a lot of different ones, and to determine which one without being an expert in classical Arabic is well beyond the scope of this paper.

Persian and Urdu texts are also difficult. Commentaries on Arabic texts are similar to the description of tafsirs above. Poetry is frequently found; it is easily recognizable by being written in couplets, usually written horizontally with a space between. As there is a lot of literature written thus it is not practical to go into details; sometimes it is possible to get something out of section headings, so that the name Iskandar Shah indicates that it is from a copy of one of the Alexander legends. If the text is in four columns, and usually in very fine Nastaliq script, it is likely to be a copy of the Shahnameh (the Persian national epic), although it might be an imitation such as the Akbarnameh (epic of the life of the Mughal emperor Akbar); reading the headings should make it easy to tell.

Sikh literature is written in Punjabi script for the most part. The colophons read like Hindu colophons. The main work is the Adi Granth, which is the main Sikh holy book¹⁹. It is large, whereas most Sikh books are much smaller.

Jain literature is voluminous, and complicated. The most commonly found text is the Kalpasutra²⁰, a compilation of Jain worthies. The manuscripts can be distinguished largely by the

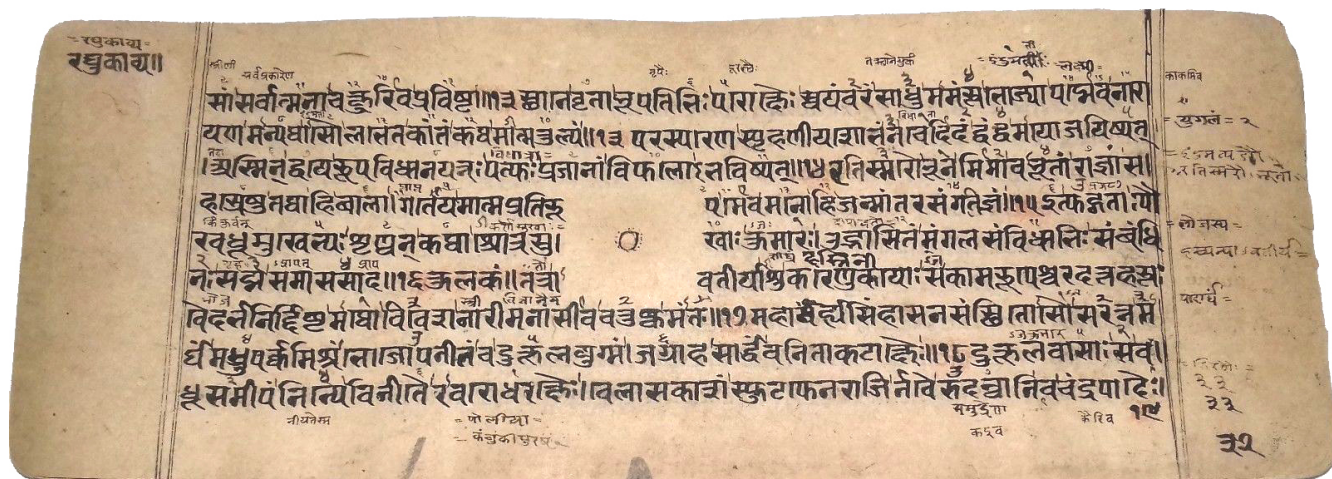
19 https://en.wikipedia.org/wiki/Guru_Granth_Sahib (3 March 20)

20 https://en.wikipedia.org/wiki/Kalpa_S%C5%ABtra (3 March 20)

decoration at the beginning of the manuscript; it resembles the number 80 in many Indian scripts. For this literature the web site “Jainpedia” covers the topic very well²¹.

Hindu literature is even more complex, as befits a literary tradition going back at least thirty-five centuries, to the Vedic period²². Vedic texts can often be distinguished by their notation, as many of them were originally chants to be sung. Most people have heard of the epics, the Mahabharata, with its subsection the Bhagavad Gita, and the Ramayana, which appear in these texts. Other Sanskrit epics appear as well, such as the epic poems of the fifth-century author Kalidasa. The title is usually found in the colophon of the manuscript, and it may be possible to find it in a reference; but the literature is vast, and largely uncatalogued.

Occasionally it is possible to determine a text from other notes; in particular, some MSS have an abbreviated title in the left margin. Sometimes it is possible to determine the text from that, with a bit of imagination and luck.



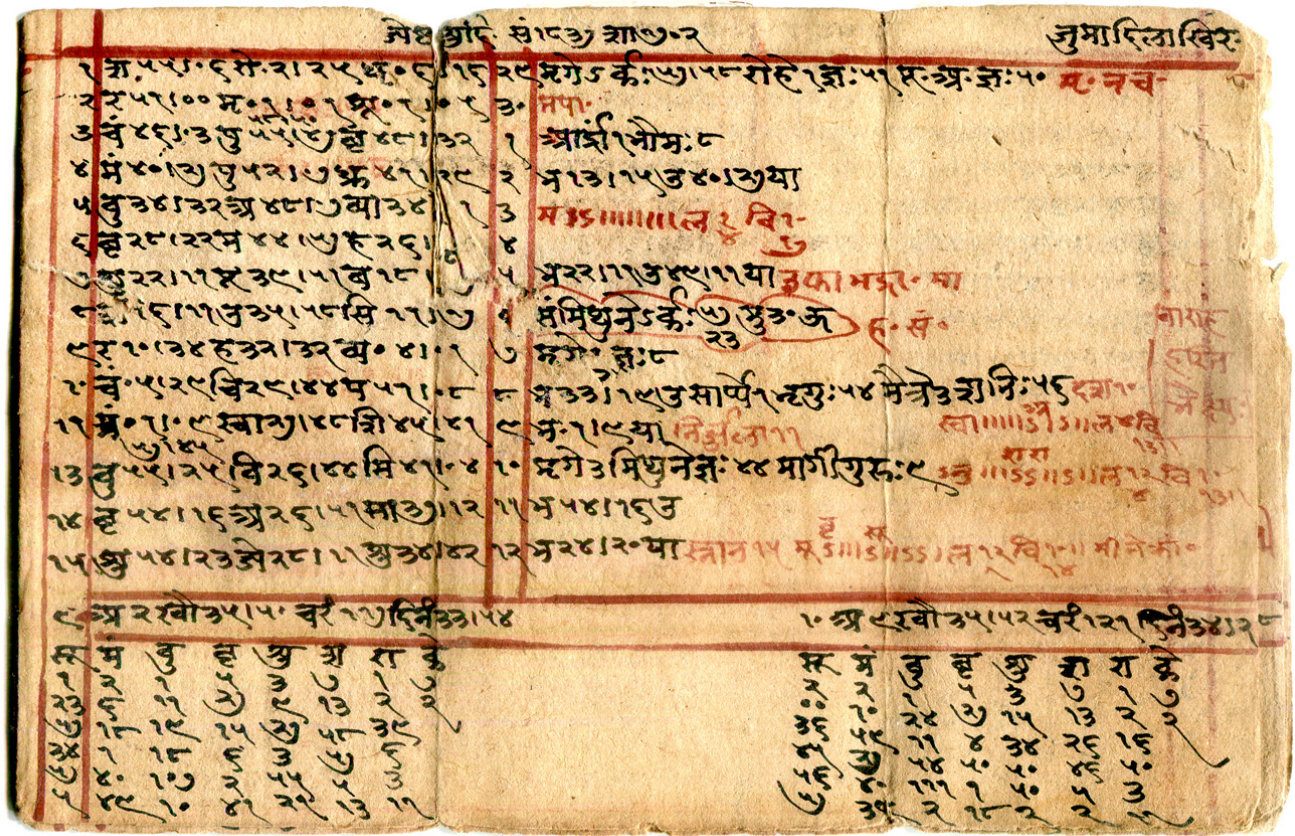
Raghuvamsa of Kalidasa, XV Century, leaf 32. The abbreviated title, Raghu kavya, is in the upper left margin.

21 <http://www.jainpedia.org/> (3 March 20)

22 https://en.wikipedia.org/wiki/Hindu_texts (3 March 20)

One easy sort of literature to spot is the almanac, which is nowadays very frequently met with.

The texts consist mostly of tables, with each page of tables having a caption at the top, usually listing the month, and year both in Samvat and Saka. The examples I have seen range on age from the mid-late eighteenth century to after 1900. Occasionally they are block-printed rather than hand-written, mostly in the 1840s and '50s.



Almanac example. The heading shows the month (Jyeshtha; May-June) and year (Samvat 1837, Saka 1702; AD 1780).

Buddhist literature is similar in terms of research, albeit with different titles. The canonical Buddhist texts are the Three Baskets (Tripitaka)²³, which are a series of ancient texts, of varied content

²³ <https://en.wikipedia.org/wiki/Tripitaka> (3 March 20)

from sermons (e. g. the “First Sermon”)²⁴, or general sayings of the Buddha (e. g. Dhammapada)²⁵, to theoretical texts. Of particular interest are the Jatakas, stories of the Buddha’s previous lives, which are a series of ancient Indian folk-tales, often with a Buddhist moral point²⁶. In addition there is the usual vast amount of commentary and miscellaneous literature, well beyond the scope of this paper.

Christian literature is fairly easy to recognize; it is either introduced from the West, on any of the local languages, or it is more or less Nestorian Syriac literature in local script, usually Malayalam, typically books of the Bible or liturgy, from the Malabar Coast²⁷.

6. Conclusions

24 https://en.wikipedia.org/wiki/Dhammacakkappavattana_Sutta (3 March 20)

25 <https://en.wikipedia.org/wiki/Dhammapada> (3 March 20)

26 https://en.wikipedia.org/wiki/Jataka_tales (3 March 20)

27 https://en.wikipedia.org/wiki/Saint_Thomas_Christians (3 March 20)