# Sanskrit Meter — Structural Reference (v1)

A foundation document for the **Sanskrit Meter** project: the metrical systems of Sanskrit described in a form precise enough to map against nucleotide and protein sequence structure. Every pattern in the tables below was generated programmatically from its gaṇa formula and machine-verified (syllable counts, algorithm round-trips, combinatorial identities), so these strings can be treated as ground truth for downstream encoding work.

**Notation used throughout:** `L` = laghu (light syllable), `G` = guru (heavy syllable), `×` = anceps (metrically free position), `|` = yati (caesura). Patterns read left to right.

---

## 1. The binary alphabet: laghu and guru

Sanskrit prosody weighs syllables, and the weighing is strictly binary.

A syllable is **guru (heavy)** if any of the following hold:

- its vowel is long (ā, ī, ū, ṝ, e, ai, o, au), or
- its (short) vowel is followed by a consonant cluster, or
- it carries anusvāra (ṃ) or visarga (ḥ).

Otherwise it is **laghu (light)**: a short vowel (a, i, u, ṛ, ḷ) followed by at most one consonant. Weight is a property of the syllable as spoken in context — a short vowel becomes heavy "by position" when the next word begins with a cluster.

One systematic relaxation: the **final syllable of a pāda (verse quarter) is anceps** — it may be light or heavy and is counted as whatever the meter needs (*pādānta* license). This is the one position where every meter loosens its grip; it returns as a structural parallel in §7.

Equivalent encodings, fixed here as project conventions:

| Symbol | Western | Bit (proposed) | Morae (mātrā) |
|---|---|---|---|
| L | ˘ | 0 | 1 |
| G | ¯ | 1 | 2 |

A pāda of *n* syllables is therefore a point in $\{L,G\}^n$ — a binary string — and a meter is a constraint on that space (one exact string for fully fixed meters, a sublanguage for freer ones). The mora value adds a second, *weighted* reading of the same string: duration rather than position.

---

## 2. The gaṇa code: reading in triplets

Classical prosody parses syllable strings in threes. A **gaṇa** is one of the $2^3 = 8$ possible triplets, each named by a single letter:

| Gaṇa | Pattern | Bits (L=0, G=1) | Value | Morae |
|---|---|---|---|---|
| ma | G G G | 111 | 7 | 6 |
| ya | L G G | 011 | 3 | 5 |
| ra | G L G | 101 | 5 | 5 |
| sa | L L G | 001 | 1 | 4 |
| ta | G G L | 110 | 6 | 5 |
| ja | L G L | 010 | 2 | 4 |
| bha | G L L | 100 | 4 | 4 |
| na | L L L | 000 | 0 | 3 |

Leftover syllables outside complete triplets are written `la` (single L) and `ga` (single G). A meter's definition is a gaṇa formula: Vasantatilakā = *ta bha ja ja ga ga*.

Two facts about this table matter for the project:

**The citation order is an enumeration algorithm.** The traditional order of the gaṇas — *ma ya ra sa ta ja bha na* — is exactly the order in which Piṅgala's prastāra procedure (§6) generates the 8 triplets. The gaṇa alphabet is literally "counting from 0 to 7 in binary," fixed in the tradition's own canonical ordering.

**The mnemonic is a de Bruijn sequence.** The traditional memory-word ***yamātārājabhānasalagāḥ*** has the weight string:

```
ya mā tā rā ja bhā na sa la gāḥ
L  G  G  G  L  G  L  L  L  G
```

Every consecutive 3-syllable window spells the gaṇa that window starts with:

| Window | Weights | Gaṇa |
|---|---|---|
| 1–3 | LGG | ya |
| 2–4 | GGG | ma |
| 3–5 | GGL | ta |
| 4–6 | GLG | ra |
| 5–7 | LGL | ja |
| 6–8 | GLL | bha |
| 7–9 | LLL | na |
| 8–10 | LLG | sa |

All 8 triplets appear exactly once in 8 overlapping windows — a binary de Bruijn sequence B(2,3), verified above, centuries before de Bruijn. The overlapping-window reading is the same formal trick as k-mer decomposition in genome assembly (§7).

---

## 3. The Vedic system: counted syllables, fixed cadence

The oldest layer (Ṛgveda onward) defines a meter by **total syllable count and pāda structure**, not by a fixed L/G value at every position. Only the cadence — the last 4–5 syllables of a pāda — settles into a preferred rhythm; the opening is quantitatively free. The canonical seven:

| Meter | Syllables | Structure | Typical cadence (per pāda) |
|---|---|---|---|
| Gāyatrī | 24 | 3 × 8 | L G L × |
| Uṣṇih | 28 | 8+8+12 | (mixed) |
| Anuṣṭubh | 32 | 4 × 8 | L G L × |
| Bṛhatī | 36 | 8+8+12+8 | (mixed) |
| Paṅkti | 40 | 5 × 8 | L G L × |
| Triṣṭubh | 44 | 4 × 11 | G L G × |
| Jagatī | 48 | 4 × 12 | G L G L × |

Structural signature of the family: **each step adds exactly 4 syllables**, and the series continues in +4 increments past the canonical seven — Atijagatī (52), Śakvarī (56), Atiśakvarī (60), Aṣṭi (64), Atyaṣṭi (68), Dhṛti (72), Atidhṛti (76), Kṛti (80), Prakṛti (84), Ākṛti (88), Vikṛti (92), Saṃkṛti (96), Abhikṛti (100), Utkṛti (104).

A disambiguation that prevents real confusion: the classical treatises reuse these same names for **classes by syllables-per-pāda** (1 = Uktā, 2 = Atyuktā, … 6 = Gāyatrī, 11 = Triṣṭubh, 19 = Atidhṛti, … 26 = Utkṛti). In that frame "Gāyatrī" means the class of all meters with 6-syllable pādas, and a specific meter like Śārdūlavikrīḍita (19 syllables/pāda) "belongs to the Atidhṛti class." Vedic usage counts the whole stanza; classical taxonomy counts the pāda. This reference states syllables **per pāda** unless marked otherwise.

The Vedic type constrains *statistics* (count + cadence) rather than *sequence* — the interior of a Vedic pāda is variable, the boundary conserved.

---

## 4. Classical varṇavṛtta: every position fixed

Classical poetry's dominant system fixes the exact weight of every syllable. Three subtypes by pāda symmetry: **samavṛtta** (all four pādas identical — the common case), **ardhasamavṛtta** (pādas 1/3 and 2/4 match), **viṣamavṛtta** (all four differ). Final syllable anceps throughout (§1).

### 4.1 The śloka (semi-fixed)

The epic meter — the bulk of the Mahābhārata and Rāmāyaṇa — is a half-fixed descendant of the Vedic anuṣṭubh: 8 syllables × 4 pādas, with only the cadence region determined. In the standard (*pathyā*) form:

```
odd pādas (1,3):   × × × ×  L G G ×
even pādas (2,4):  × × × ×  L G L ×
```

with the added rule that syllables 2–3 of any pāda may not both be L. Named *vipulā* variants substitute other cadences (na-, bha-, ma-, ra-vipulā) in odd pādas under stated conditions. The śloka is thus a *constrained sublanguage* of $\{L,G\}^8$ rather than a single string — the most protein-like of the classical meters in that respect.

### 4.2 Samavṛtta catalog (fully fixed)

Machine-verified patterns, one pāda each (× 4 for the stanza). Yati positions mark obligatory word-break groupings.

| Meter | Syll/pāda | Gaṇa formula | Pattern (per pāda) | Yati |
|---|---|---|---|---|
| Indravajrā | 11 | ta ta ja ga ga | GGLGGLLGLGG | — |
| Upendravajrā | 11 | ja ta ja ga ga | LGLGGLLGLGG | — |
| Upajāti | 11 | mix of the two above | ×GLGGLLGLGG | — |
| Rathoddhatā | 11 | ra na ra la ga | GLGLLLGLGLG | — |
| Śālinī | 11 | ma ta ta ga ga | GGGGGLGGLGG | 4+7 |
| Vaṃśastha | 12 | ja ta ja ra | LGLGGLLGLGLG | — |
| Indravaṃśā | 12 | ta ta ja ra | GGLGGLLGLGLG | — |
| Bhujaṅgaprayāta | 12 | ya ya ya ya | LGGLGGLGGLGG | — |
| Toṭaka | 12 | sa sa sa sa | LLGLLGLLGLLG | — |
| Drutavilambita | 12 | na bha bha ra | LLLGLLGLLGLG | — |
| Sragviṇī | 12 | ra ra ra ra | GLGGLGGLGGLG | — |
| Praharṣiṇī | 13 | ma na ja ra ga | GGGLLLLGLGLGG | 3+10 |
| Rucirā | 13 | ja bha sa ja ga | LGLGLLLLGLGLG | 4+9 |
| Vasantatilakā | 14 | ta bha ja ja ga ga | GGLGLLLGLLGLGG | — |
| Mālinī | 15 | na na ma ya ya | LLLLLLGGGLGGLGG | 8+7 |
| Pañcacāmara | 16 | ja ra ja ra ja ga | LGLGLGLGLGLGLGLG | — |
| Śikhariṇī | 17 | ya ma na sa bha la ga | LGGGGGLLLLLGGLLLG | 6+11 |
| Mandākrāntā | 17 | ma bha na ta ta ga ga | GGGGLLLLLGGLGGLGG | 4+6+7 |
| Hariṇī | 17 | na sa ma ra sa la ga | LLLLLGGGGGLGLLGLG | 6+4+7 |
| Pṛthvī | 17 | ja sa ja sa ya la ga | LGLLLGLGLLLGLGGLG | 8+9 |

| Meter | Syll/pāda | Gaṇa formula | Pattern (per pāda) | Yati |
|---|---|---|---|---|
| Śārdūlavikrīḍita | 19 | ma sa ja sa ta ta ga | GGGLLGLGLLLGGGLGGLG | 12+7 |
| Sragdharā | 21 | ma ra bha na ya ya ya | GGGGLGGLLLLLLGGLGGLGG | 7+7+7 |

Notes worth keeping: Pañcacāmara is perfectly periodic — $(LG)^8$ — the maximally alternating string. Mandākrāntā's yati grouping `GGGG | LLLLLG | GLGGLGG` moves from all-heavy through all-light to mixed, an engineered rhythmic gradient (Kālidāsa's *Meghadūta* meter). Meters longer than 26 syllables per pāda are collectively **daṇḍaka**.

### 4.3 Ardhasamavṛtta and viṣamavṛtta

| Meter | Pādas | Patterns |
|---|---|---|
| Viyoginī | odd 10 / even 11 | LLGLLGLGLG / LLGGLLGLGLG |
| Puṣpitāgrā | odd 12 / even 13 | LLLLLLGLGLGG / LLLLGLLGLGLGG |
| Udgatā (viṣama) | 10, 10, 11, 13 | LLGLGLLLGL / LLLLLGLGLG / GLLLLLLGLLG / LLGLGLLLGLGLG |

---

## 5. Mātrā meters: fixed duration, variable sequence

The third system counts **morae** (mātrās: L = 1, G = 2) and fixes total duration per unit while leaving the L/G realization open — fixed "mass," many admissible sequences.

**The Āryā family** is the center of gravity: a stanza of two halves of **30 + 27 mātrās**, each half organized as 7½ four-mora gaṇas (the half-gaṇa is a final 2-mora unit; in the second half the 6th gaṇa is a lone laghu, which is what shortens 30 to 27). There are exactly **five** four-mora patterns available as building blocks:

```
GG   LLG   LGL   GLL   LLLL
```

with placement rules (the amphibrach LGL is barred from odd gaṇa positions; the 6th gaṇa of the first half must be LGL or LLLL). Variants by half-combinations: Gīti (30+30), Upagīti (27+27), Udgīti (27+30), Āryāgīti (32+32).

**Vaitālīya**: odd pādas 14 mātrās, even pādas 16, each closing with the fixed 8-mora cadence G L G L G. Related: Aupacchandasika, Mātrāsamaka.

The combinatorics of this system is the historically celebrated one: the number of L/G sequences totaling *m* morae is the **Virahāṅka number** — 1, 2, 3, 5, 8, 13, 21, 34, … for m = 1, 2, 3, … — the Fibonacci sequence, derived by Virahāṅka (c. 7th c. CE) and successors from exactly this metrical question, five centuries before Fibonacci. (Verified above; note the five 4-mora gaṇas are the m = 4 case.) Mātrā space is the space of **compositions of an integer into parts {1, 2}**.

---

## 6. Piṅgala's algorithms: the six pratyayas

Piṅgala's *Chandaḥśāstra* (traditionally c. 3rd century BCE; dating debated) closes with procedures (*pratyayas*) that treat meter-space as a formal combinatorial object. With L = 0 / G = 1 these are recognizably binary algorithms.

**Prastāra (enumeration).** Generate all $2^n$ patterns of n syllables in canonical order: start with all-G; at each step find the leftmost G, change it to L, set everything to its left to G, copy the rest. For n = 3 this yields rows 1–8 = GGG, LGG, GLG, LLG, GGL, LGL, GLL, LLL — i.e. *ma ya ra sa ta ja bha na*, the traditional gaṇa order (§2). The row index k−1, written in binary least-significant-bit-first with L = 1, G = 0, reproduces the row: prastāra **is** binary counting.

**Naṣṭa ("lost" — index → pattern).** To recover row k without enumerating: repeatedly, if k is odd write G and set k ← (k+1)/2; if even write L and set k ← k/2; n steps, left to right. (Row 5, n=3: 5→G,3→G,2→L = GGL = ta.)

**Uddiṣṭa ("indicated" — pattern → index).** Inverse: start v = 1; scan the pattern right to left; for L set v ← 2v, for G set v ← 2v−1. The two procedures were machine-verified as mutual inverses over all rows for n ≤ 8. Worked example: Vasantatilakā GGLGLLLGLLGLGG is **row 2933 of the 16384** patterns of length 14.

**Saṃkhyā (count).** Total patterns = $2^n$, computed by Piṅgala with a halve-and-square recursion — "if even, halve and square; if odd, subtract one and double" — i.e. **exponentiation by squaring**, the earliest known statement of it.

**Lagakriyā (count by composition).** How many n-syllable patterns contain exactly k gurus: the binomial coefficient C(n,k), tabulated by the commentator Halāyudha (10th c.) as the **meru-prastāra** — Pascal's triangle, seven centuries before Pascal.

**Adhvayoga (total space).** Patterns of all lengths 1…n: $2^{n+1} - 2$. For the classical taxonomy's 26 pāda-lengths: $2^{27} - 2$ = **134,217,726** possible pāda-patterns, of which the named, used meters are a few hundred — the functional subset is on the order of $10^{-6}$ of sequence space.

For mātrā meters the tradition developed parallel pratyayas (mātrā-prastāra etc.) whose counting function is the Virahāṅka–Fibonacci recurrence of §5.

---

## 7. Mapping hooks: meter ↔ DNA / protein

The load-bearing parallels, graded by how much weight they can carry. **Exact** = mathematical identity; **structural analogy** = same organizing principle, different substrate; **caution** = popular but not load-bearing.

| # | Meter side | Molecular side | Status |
|---|---|---|---|
| 1 | Binary alphabet {L,G}, 1 bit/ syllable | Binary *reductions* of {A,C,G,T/U}: purine/ pyrimidine R/Y, strong/weak S/W, amino/keto M/K — each 1 bit/base of the 2-bit alphabet | **Exact** (as projection) |
| 2 | Triplet parsing: $2^3 = 8$ gaṇas | Triplet codons: $4^3 = 64$; under any binary reduction (e.g. R/Y) the 64 codons collapse to exactly **8 classes of 8** — isomorphic to gaṇa space | **Exact** |
| 3 | De Bruijn mnemonic (§2): all triplets via overlapping windows | k-mer / de Bruijn-graph decomposition in genome assembly; overlapping reading frames | **Exact** (same formalism) |
| 4 | Pāda-final anceps: the one systematically free position (§1) | Codon third-position wobble: the systematically degenerate position | Structural analogy (unit-boundary relaxation in both codes) |
| 5 | Vedic type: free interior + conserved cadence (§3) | Variable regions + conserved motifs at functional sites | Structural analogy |
| 6 | Named meters ≈ $10^{-6}$ of the $1.3 \times 10^8$ pāda space (§6) | Functional proteins as a vanishing subset of sequence space | Structural analogy |
| 7 | Mātrā meters: fixed total duration, variable sequence; count = Fibonacci compositions into {1,2} (§5) | Fixed-property-sum classes over variable sequences (e.g. mass classes); {1,2}-composition combinatorics | Structural analogy |
| 8 | Śloka: constrained sublanguage, not a single string (§4.1) | Sequence families / motifs (PWMs) rather than exact consensus | Structural analogy |
| 9 | Gāyatrī's 24 ↔ various molecular 24s; chakra/deity numerologies | — | **Caution**: numerological; treat as cultural data, not structure |

Row 2 is the cleanest bridge and suggests the project's first concrete artifact: a **codon → gaṇa dictionary**. It requires one choice — which binary reduction, and which polarity. Two natural candidates: R→G / Y→L (purines are the larger bases → "heavy"), or S→G / W→L (G≡C pairs bind with three hydrogen bonds vs two → "heavy" as *stronger/longer*, which echoes guru as *longer duration*). The S/W choice is the more physically motivated analogy; both should be tested rather than assumed. Under R→G/Y→L, for the record: RRR→ma, YRR→ya, RYR→ra, YYR→sa, RRY→ta, YRY→ja, RYY→bha, YYY→na.

Honesty constraints for the project: the mapping in row 2 is exact only as a *projection* (a gaṇa corresponds to an equivalence class of 8 codons, not a codon); rows 4–8 share organizing principles, not mechanisms; and any claim routed through row 9 should be flagged as such.

---

## 8. Proposed data schema

For the machine-readable catalog (next artifact), one record per meter:

```
{
  "name_iast": "Vasantatilakā",
  "system": "varnavrtta",            // vedic | varnavrtta | ardhasama | vishama | matra
  "class_by_pada": "Śakvarī (14)",
  "syllables_per_pada": 14,
  "ganas": "ta bha ja ja ga ga",
  "pattern_LG": "GGLGLLLGLLGLGG",     // per pāda; array for non-sama systems
  "bits_G1": "11010001001011",        // L=0, G=1
  "prastara_row": 2933,               // uddiṣṭa index within its 2^n class
  "morae_per_pada": 21,
  "yati": null,
  "final_anceps": true,
  "notes": ""
}
```

Conventions fixed by this doc: G = 1, L = 0 for bits; morae L = 1, G = 2; prastāra row via uddiṣṭa (§6); patterns stated per pāda with final position recorded as written but flagged anceps.

---

## 9. Sources

Primary treatises: Piṅgala, *Chandaḥśāstra*, with Halāyudha's commentary *Mṛtasañjīvanī*; Kedārabhaṭṭa, *Vṛttaratnākara* (the standard classical catalog; source of the gaṇa-formula definitions used in §4); Gaṅgādāsa, *Chandomañjarī*.

Modern scholarship: H. D. Velankar, *Jayadāman* (1949) — critical edition of four chandas treatises with the standard index of meters; E. V. Arnold, *Vedic Metre in its Historical Development* (1905) — for §3; A. A. Macdonell, *A Vedic Grammar for Students*, appendix on meter.

Mathematics: B. van Nooten, "Binary numbers in Indian antiquity," *Journal of Indian Philosophy* 21 (1993) — prastāra as binary arithmetic; Parmanand Singh, "The so-called Fibonacci numbers in ancient and medieval India," *Historia Mathematica* 12 (1985) — Virahāṅka; D. E. Knuth, *TAOCP* vol. 4A, §7.2.1.1 — credits the gaṇa mnemonic as the oldest known de Bruijn cycle and Piṅgala's pratyayas as early combinatorial generation.

*Patterns and identities in this document were regenerated and checked by script (verify_meters.py in the session workspace); the doc-table strings were diffed against the generated set before delivery.*