

数字渐开线简介——渐开线反函数的离散归正法

陈典章

chendianzhang@163.com

摘要

传统反渐开线函数求解依赖迭代或级数展开，面临初值敏感、分段缝合、多精度仿真无法在有限字长硬件上物理复现等困境。本文给出数字渐开线：以曲率域渐开线方程 $1/\tan a = 1/(k + a)$ 为数学本体，以归正方程 $U(1 + UK) = D$ 为工程形式，在 IEEE 754 硬件上完成归正的投影过程——离散，终态即硬件最小不可分辨的离散坐标。

给出固定顺序的初等运算链（双精度工程链与四精度基准链），用于在给定展角 $k = \tan a - a$ 下直接生成渐开线齿廓的极坐标 (θ, ρ) 。该链不含分支、迭代、截断或人为阈值，无外部参考，全域适用，时延固定（一次 $\arctan$ 、一次 $\tan$ ），精度仅由浮点数舍入误差决定。前向代入检验 $1/\tan a \ominus 1/(k + a) = 0$ 由硬件零标志位锁定终态唯一。

双精度下误差谱表明：归真输出 a 始终锁定在硬件邻域内，在 $k > 10^{-4}$ 时参考误差进入 10^{-15}rad 量级，深零域邻域半径峰值约 $2.1 \times 10^{-13}\text{rad}$ ；增强速算 W 在 $k < 0.36$ ($\alpha < 52^\circ$) 精度达硬件极限，相对于高精度参考真值的全域最大误差约 $6 \times 10^{-12}\text{rad}$ ；轻量速算 M 在 $k < 15$ ($\alpha < 86^\circ$) 优于 $5 \times 10^{-9}\text{rad}$ ，覆盖工控需求。

数字渐开线是理论渐开线在浮点网格上的离散投影点列，可作为齿轮设计、制造、计量的数字基准路径，不依赖物理展成链。

关键词：数字渐开线，曲率驱动，归正方程，硬件投影，离散，数字基准，数控

中图分类号：TH132.413；O241；TB11

The Digital Involute: A Discrete Alignment Method for the Inverse Involute Function

Dianzhang Chen

Abstract

Conventional methods for solving the inverse involute function rely on iteration or series expansion, suffering from initial-value sensitivity, piecewise stitching, and multi-precision simulation that cannot be physically reproduced on finite-word-length hardware. This paper presents the digital involute: taking the curvature-domain involute equation $1/\tan a = 1/(k + a)$ as the mathematical ontology and the alignment equation $U(1 + UK) = D$ as the engineering form, the alignment projection onto IEEE 754 hardware—discretization—is completed, with the terminal state being the smallest hardware-resolvable discrete coordinate.

A fixed-sequence elementary chain (double-precision engineering chain and quad-precision reference chain) is given for directly generating the polar coordinates (θ, ρ) of an involute tooth profile for a given roll angle $k = \tan a - a$. The chain involves no branching, iteration, truncation, or artificial threshold, requires no external reference, applies globally, has deterministic latency (one $\arctan$, one $\tan$), and its precision is bounded solely by floating-point rounding. The forward-substitution test $1/\tan a \ominus 1/(k + a) = 0$ locks the unique terminal state via the hardware zero flag.

Under double precision, error spectra show: the true-alignment output a is always locked within the hardware neighborhood, with the reference error entering the 10^{-15}rad level for $k > 10^{-4}$ and a deep-zero-domain neighborhood radius peak of approximately $2.1 \times 10^{-13}\text{rad}$; the enhanced rapid-calculation W reaches the hardware limit for $k < 0.36$ ($\alpha < 52^\circ$), with a global maximum error relative to the high-

precision reference truth of approximately $6 \times 10^{-12} \text{rad}$; the lightweight rapid-calculation M is better than $5 \times 10^{-9} \text{rad}$ for $k < 15 (\alpha < 86^\circ)$, covering industrial control requirements.

The digital involute is a discrete projection point sequence of the theoretical involute onto the floating-point grid, and can serve as a digital reference path for gear design, manufacturing, and metrology, independent of physical generating chains.

Keywords: digital involute, curvature-driven, alignment equation, hardware projection, discretization, digital reference, CNC

0. 引言

传统反渐开线函数求解方法，无论是迭代法还是级数展开法，均将求解视为对连续真值的无限逼近过程。迭代法依赖初值选择与人为收敛判据，步数不定，在零域与极域易震荡、发散；级数法则受限收敛半径，必须分段拼接，在缝合处引入导数断裂与精度塌陷，模型定义多值则唯一性失真。多精度仿真虽可生成高位数结果，却无法在有限字长硬件上物理复现。这些方法共同的根源在于：将精确性寄托于外部参考与人为阈值，而非硬件自身的离散判定。

数字渐开线以曲率域渐开线方程 $1/\tan a = 1/(k + a)$ 为数学本体，以归正方程 $U(1 + UK) = D$ 为工程形式，在 IEEE 754 硬件上完成归正的投影过程——离散。偏置、归正、确认、锁定四环节构成固定算术序列，终态即硬件最小不可分辨的离散坐标。这就是数字渐开线的工程极限定义。

1. 数字渐开线链

给定展角 $k = \tan a - a$ ，齿廓点极坐标：

$$\theta = k, \quad \rho = r_b \sqrt{1 + (k + a)^2}, \quad (1)$$

压力角 a 由以下链式确定，式 (3) 适配双精度工程链输出。

$$C = (0.6 + 0.08\sqrt{k})k + \sqrt[3]{3k}, \quad G = \arctan C, \quad P = k + G, \quad D = 1/C - 1/P, \quad (2)$$

$$A = G + D - \frac{D^2}{P}, \quad a = A + \frac{1}{\tan A} - \frac{1}{k + A}, \quad (3)$$

记 $A = G + M$ 为轻量速算输出，其中 $M = D - D^2/P$ ，对应双精度工程链。该链全域 $k > 0$ 无分段，时延固定，精度由硬件浮点数舍入直接限定，无迭代逼近、级数截断或人为判据。

归真输出 a (计量基准)：双精度下全域精度不低于 $2.1 \times 10^{-13} \text{rad}$ (邻域半径峰值在深零域 $k \approx 2 \times 10^{-8}$)，在 $k > 10^{-4}$ 零域以外不低于 $2 \times 10^{-15} \text{rad}$ 。

速算输出 A (实时控制，取 $a \approx A$)：全域精度不低于 $2 \times 10^{-8} \text{rad}$ ；在 $10^{-6} < k < 0.36$ 常用域内不低于 $2.17 \times 10^{-11} \text{rad}$ 。单精度 FPU 速算 A 精度约 10^{-6}rad 量级。

附注 (验证与扩展精度)

对 binary128 FPU (IEEE 754) 硬件或 mpmath 仿真环境，式 (4) 适配四精度基准链输出：

$$A = G + \frac{D}{3} \left(\frac{4}{1 + \sqrt{1 + 4D/P}} + \frac{1}{1 + D/C} \right), \quad d = \frac{1}{\tan A} - \frac{1}{k + A}, \quad a = A + \frac{d}{1 + d/(k + A)}, \quad (4)$$

记 $A = G + W$ 为增强速算输出，其中 $W = \frac{D}{3} \left(\frac{4}{1 + \sqrt{1 + 4D/P}} + \frac{1}{1 + D/C} \right)$ ，对应四精度基准链。

归真输出 a ：四精度输出齿形精度约 10^{-33}rad 量级。

在双精度下，该链归真输出 a 精度与工程链 (3) 式结果一致。逐点速算 (取 $a \approx A$ 时) 输出 A 全域精度不低于 6×10^{-12} (在 $10^{-6} < k < 0.36$ 常用域内不低于 8×10^{-15})，可作为工程计量的数字基准。若密步进到新 k 连续归真，用当前点归真 a 赋值下一点速算 A ，按 $a \Rightarrow A \rightarrow d \rightarrow a$ 递推。

2. 离散归正原理

连续域方程 $\tan a - a - k = 0$ 在开定义域 $(0, \pi/2)$ 上的真值为无理数，在浮点网格 $\mathbb{F}$ 上无法连续解析表达。离散归正将其重构为：在 $\mathbb{F}$ 上寻找唯一满足曲率对齐的有理数格点，通过偏置 $\rightarrow$ 归正 $\rightarrow$ 确认链锁定。其终态由前向代入检验 $1/\tan a \ominus 1/(k+a) = 0$ 的 IEEE 754 硬件零标志位确定，归真输出即为给定字长下唯一可复现的有理投影点列。

偏置：全域连续光滑的曲率半径引导函数 $C = 0.68k + \sqrt[3]{3k} + 0.08(\sqrt{k} - 1)k$ ，粗率 $1/C$ ，投影角 $G = \arctan C$ ，构造曲率半径 $P = k + G$ ，曲率残差 $D = 1/C - 1/P$ ，全局线性有界。构造系数为 $0.68 = (0.6 + 0.08) \approx 2/3$ ，用于曲率偏置与归正路径的归一调制优化。

归正：基本归正量 U ，速算角 $A = G + U$ ，精率 $1/c = 1/\tan A$ ，构造精率半径 $p = k + A$ 。

展开几何约束 $1/\tan(G+U) = 1/(k+G+U)$ ，得曲率残差 D 驱动的二阶微量归正方程：

$$U = \frac{D}{1 + U/P} - R^+ \left(1 + \frac{1}{C \tan A}\right), \quad R^+ = \tan U - U = O(U^3)$$

略去三阶微量 R^+ ，得工程等价形式： $U + U^2/P \simeq D$ ，其有理解为： $U = D/(1 + D/P)$ 。

$$\text{基本归正 } \frac{1}{U} - \frac{1}{D} = \frac{1}{P}, \quad \text{调制路径 } \frac{1}{U} - \frac{1}{D} = K \simeq \frac{1}{T}$$

调制曲率 $K = \lambda/P + (1-\lambda)/C$ ，调制曲率半径 $T = \lambda P + (1-\lambda)C$ ， $\lambda \in [0,1]$ 权重构造路径：

$$\begin{aligned} T_U &= P, \quad T_M = P - D, \quad T_V = P(1 + \sqrt{1 + 4D/P})/2 \approx P + M, \quad T_0 = C. \\ U &\simeq \frac{D}{1 + DK} \simeq \frac{D}{1 + D/T} \simeq D \left[\frac{\lambda}{1 + D/P} + \frac{1-\lambda}{1 + D/C} \right], \quad V = \frac{D}{1 + D/T_V} \approx D \left[\frac{\lambda}{1 + D/T_V} + \frac{1-\lambda}{1 + D/C} \right], \\ U &\simeq \frac{D}{1 + D/P} \approx \text{轻量 } M = D \left(1 - \frac{D}{P}\right), \quad \text{增强 } W = \frac{D}{3} \left[\frac{4}{1 + \sqrt{1 + 4D/P}} + \frac{1}{1 + D/C} \right] = M + O(D^3). \end{aligned}$$

确认：基本归真量 u ，归真角 $a = A + u$ 。真率 $1/\tan a$ ，构造真率 $1/(k+a)$ 。

精率残差 $d = 1/\tan A - 1/(k+A)$ ，直接归真 d 用于工程链轻量。

$$\text{基本归真 } \frac{1}{u} - \frac{1}{d} = \frac{1}{p}, \quad u = \frac{d}{1 + d/p} = d + O(d^2)$$

终态判据：真率残差 $1/\tan a - 1/(k+a) \lesssim \varepsilon_{mach}$

锁定：邻域格点 $\delta p/p = P\delta a = \delta k/P \sim ulp$ 。

邻域半径 δa 由硬件字长与渐开线灵敏度 $P = \tan a$ 共同界定，在 ulp 尺度下终态不可分辨。

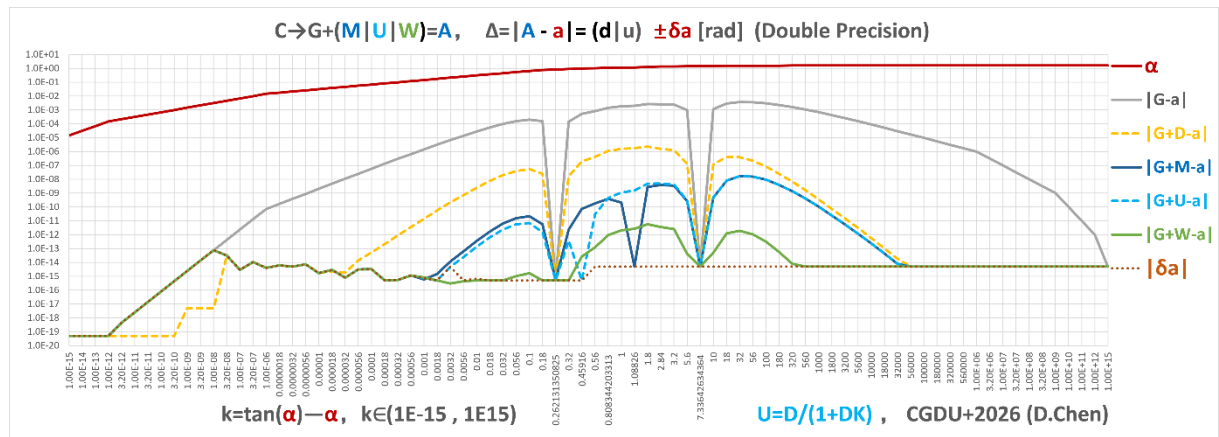


图1 归正速算残留偏离分布与归真确认邻域半径的全域图谱（横轴以展角 k 为对数尺度）

横轴为展角 k (对数)，纵轴为双精度下归正角度偏差 Δ 与归真邻域半径 δa (对数)，典型误差界：

增强速算 $(A = G + W) \sim$ 有理速算 $(A = G + U) \approx$ 轻量速算 $(A = G + M) \approx$ 直接速算 $(A = G + D)$ 。

由于双精度 (binary64) 硬件可提供的最大安全相对精度约为 15 位十进制有效数字 ($10^{-15} > \varepsilon_{\text{mach}} > 10^{-16}$)，数字渐开线的有效输入展角全域取为 $k \in [10^{-15}, 10^{15}]$ 。在此范围内，终态锁定的极径输出 ρ 为最小可分辨的相邻 ULP 格点；确认角度输出 α 落在硬件邻域内，真率残差归零。

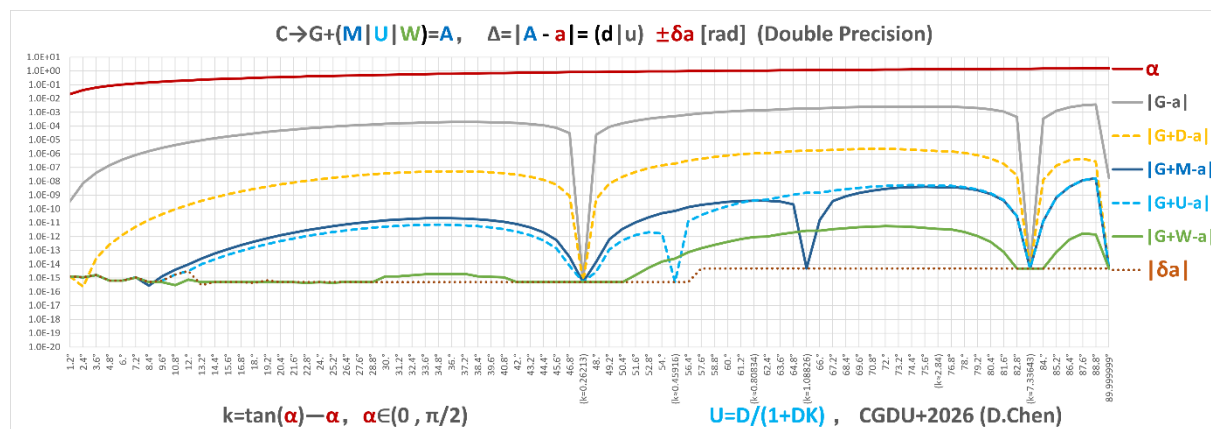


图 2 归正速算残留偏离分布与归真确认邻域半径的全域图谱（横轴取角度 α 为线性尺度）

图中误差曲线为与高精度参考真值的偏差，参考真值由零域级数、中域迭代、极域渐近展开等方法生成； δa 为归真输出角度的邻域半径，与参考误差量级相当，在全域扫描中得到验证。

3. 结语

数字渐开线是理论渐开线在给定浮点系统下的唯一可确认投影。它以展角 k 为输入，经由偏置 $\rightarrow$ 归正 $\rightarrow$ 确认链锁定终态，输出压力角 α 与齿形极坐标 (θ, ρ) 。归正方程与渐开线方程终态等价，由 IEEE 754 硬件零标志位完成物理确认。其生成与确认由三项原则支撑：确认准则以硬件分辨率终止判据，等价原则保证归正方程与渐开线方程终态一致，映射规则锁定输出格点唯一。

双精度工程链与四精度基准链共用同一代数底座：前者面向当下工业交付，归真输出约 10^{-15}rad 量级；后者面向未来硬件储备，可延伸至 10^{-33}rad 量级。制造坐标与计量基准共享同一点列，路径即基准。该路径 $k \rightarrow C \rightarrow G \rightarrow D \rightarrow U|M|W \rightarrow A \rightarrow c \rightarrow d \rightarrow u \rightarrow a \rightarrow \rho$ 简称为 CGDU+ 归正法，是对连续渐开线定义 $\alpha \rightarrow \tan \alpha \rightarrow k$ 的反函数重构的确定性工程路径，亦即离散域中数字渐开线的生成方式。

它不是对无理数的无限逼近，而是对离散有理格点的唯一锁定，是数字制造时代的工程极限定义。该方法使反渐开线函数求解从依赖外部判据的数值无限逼近，转变为硬件分辨率锚定的曲率对齐确认，为齿轮设计、加工、计量以及传统物理传动链的升级改造、高精度大展角齿形的精密加工与凸轮应用，提供了一条不依赖物理展成链的数字基准路径。

说明

预印本，公开即构成现有技术。本文为《工程渐开线 (CGDU+)》的公开简介，核心方法已获著作权登记（登记号：2026-A-00071724）。完整文稿登记受理中（受理号：2026Z11S0186755）。

参考文献

- [1] 张希康.(译) 求解渐开线反函数的公式[J]. 齿轮, 1980(1): 50, (摘自 ISO/TC60.413(1978).)
- [2] 陆炽华. 关于渐开线压力角的计算[J]. 工具技术, 1986(3):15-16
- [3] 周汝忠. 反渐开线函数的精确计算方法[J]. 机床, 1990(4):18-19

- [4] 陈典章. 量球(柱)测量齿厚的计算[J]. *设备管理与维修*, 1997(8): 30-33.
- [5] H. H. Cheng, "Derivation of the explicit solution of the inverse involute function and its applications in gear tooth geometry calculations," *Journal of Applied Mechanisms & Robotics*, vol. 3, no. 2, 1996.
- [6] S. Rosado, F. P. Muñoz, and R. A. Fernández, "An analytic expression for the inverse involute," *Mathematical Problems in Engineering*, vol. 2019, Article ID 3586012, 2019.
- [7] Y. Zhou, B. Wu, Z. Chen, C. Zhong, and H. Zhong, "Calculation of the inverse involute function and application to measurement over pins," *Measurement Science and Technology*, vol. 35, no. 11, p. 117002, 2024.
- [8] IEEE, "IEEE Standard for Floating-Point Arithmetic," *IEEE Std 754-2019 (Revision of IEEE 754-2008)*, pp. 1-84, Jul. 2019, doi: 10.1109/IEEE STD.2019.8766229.

作者简介: 陈典章(1963-), 男, 湖南石门人, 中国船舶 388 厂, 研究员级高级工程师, 主要研究机械工程与光电装备技术. E-mail: chendianzhang@163.com
