

Optimal Multiplicative Hashing via the Three-Distance Theorem

A Formal Proof of Knuth's Golden Ratio Recommendation

Jetxel Alfredo

Independent Researcher

Scalar Resonance Research Program, March 2026

Second edition, September 2026

Abstract. A multiplicative hash function $h(k) = \lfloor m \cdot \text{frac}(k\alpha) \rfloor$ places the keys $k = 1, 2, 3, \dots$ at the points $k\alpha \bmod 1$, and the Three-Distance Theorem (Sós 1958) says those points cut the circle into gaps of at most three lengths. We prove that for $\alpha = 1/\varphi$, where $\varphi = (1 + \sqrt{5})/2$ is the golden ratio, the largest gap is always exactly φ or φ^2 times the smallest, for every number of keys from two on. Among the multipliers with periodic continued fractions $[0; a, a, a, \dots]$ this is the smallest possible spread, and the noble numbers, whose continued fractions end in ones, are the multipliers that share it in the long run. This is the uniformity behind Knuth's recommendation of $1/\varphi$. We also measure bucket loads over 169 prime table sizes with 1000 keys per bucket. There the noble multipliers, $1/\varphi$ among them, have the lowest mean overload, but the worst case, 0.3 percent above the average, is shared by $\sqrt{2} - 1$. Multipliers with large partial quotients do much worse: $1/\pi$ reaches 7.1 percent, and its load varies 39 times more across table sizes.

Index Terms. Continued fractions, Diophantine approximation, golden ratio, hash functions, multiplicative hashing, noble numbers, Three-Distance Theorem, worst approximability

Second edition. The first edition remains available, and the corrections are listed at the end.

1. Introduction

Multiplicative hashing is among the most widely deployed families of hash functions. A key k goes to bucket $h(k) = \lfloor m \cdot \text{frac}(k\alpha) \rfloor$ of a table with m buckets, for an irrational multiplier $\alpha \in (0, 1)$. How evenly the keys spread depends only on α .

In *The Art of Computer Programming*, Volume 3, Knuth [1] recommends $\alpha = (\sqrt{5} - 1)/2 = 1/\varphi \approx 0.6180$, observing that it produces particularly uniform distributions. Practitioners have followed the recommendation for fifty years.

This paper makes the uniformity precise. The keys land at the points $k\alpha \bmod 1$ on a circle, and by the Three-Distance Theorem the first N of them cut it into gaps of at most three different lengths. We measure evenness by the ratio of the largest gap to the smallest. For $1/\varphi$ that ratio is always φ or φ^2 , and no multiplier does better in the long run (Section 3). Section 4 turns to bucket loads, where the picture is more modest: $1/\varphi$ is among the best multipliers, level with other multipliers whose partial quotients stay small.

2. Preliminaries

2.1 Continued Fractions and Convergents

Every irrational $\alpha \in (0, 1)$ has a unique continued fraction expansion $\alpha = [0; a_1, a_2, a_3, \dots]$ with positive integer partial quotients a_n . The convergent denominators satisfy $q_n = a_n q_{n-1} + q_{n-2}$, with $q_0 = 1$ and $q_1 = a_1$. The golden ratio has $\varphi = [1; 1, 1, 1, \dots]$, so $1/\varphi = [0; 1, 1, 1, \dots]$ and its convergent denominators are the Fibonacci numbers 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, $\dots$. Since every partial quotient is at least 1, no denominators grow more slowly.

2.2 The Three-Distance Theorem

The Three-Distance Theorem (Sós [2], Surányi, Świerczkowski, 1958) states that the N points $\{\alpha, 2\alpha, \dots, N\alpha\} \bmod 1$ divide the circle into gaps of at most three distinct lengths. When N is a convergent denominator q_n , only two lengths occur: $\|q_{n-1}\alpha\|$ and $\|q_n\alpha\|$, where $\|x\|$ is the distance from x to the nearest integer. The largest gap at $N = q_n$ is therefore

$$G(\alpha, q_n) = \|q_{n-1}\alpha\| + \|q_n\alpha\|.$$

Between consecutive denominators, new points split the largest gaps one at a time. A large partial quotient a_{n+1} means many points must arrive before the largest gaps are all split, which shows up as a long plateau in the largest gap (Figure 2).

2.3 Hurwitz's Theorem

Hurwitz's theorem [3] states that every irrational α has infinitely many rational approximations p/q with $|\alpha - p/q| < 1/(\sqrt{5}q^2)$. The constant $\sqrt{5}$ cannot be improved for numbers equivalent to φ under $GL_2(\mathbb{Z})$, and it can be improved for every other irrational. In this sense φ and its equivalents are the hardest numbers to approximate.

2.4 Noble Numbers

A noble number is an irrational whose continued fraction ends in $1, 1, 1, \dots$; the nobles are exactly the numbers equivalent to φ under $GL_2(\mathbb{Z})$, and they share its asymptotic Diophantine properties. The leading partial quotients affect behavior at small scales only; the tail decides the long run.

3. The Gap Ratio

For $N \geq 2$ let $R(\alpha, N)$ be the ratio of the largest to the smallest of the gaps cut by the points $k\alpha \bmod 1, 1 \leq k \leq N$. It does not depend on scale, and it is 1 only for a perfectly even partition.

Theorem 1. For $\alpha = 1/\varphi$ and every $N \geq 2$, $R(\alpha, N)$ is either φ or φ^2 .

Proof. For $\alpha = 1/\varphi$ every partial quotient is 1, and the distances $\|q_n\alpha\|$ are consecutive powers of $1/\varphi$, each the sum of the next two. By the Three-Distance Theorem, for $q_n \leq N < q_{n+1}$ the gaps have lengths among $\|q_n\alpha\|$, $\|q_{n-1}\alpha\|$ and their sum $\|q_{n-2}\alpha\|$. At $N = q_n$ only the two longer ones are present, and their ratio is φ . Each point added after that falls in a gap of length $\|q_{n-2}\alpha\|$

and splits it into gaps of lengths $\|q_{n-1}\alpha\|$ and $\|q_n\alpha\|$, in the golden ratio, so for $q_n < N < q_{n+1}$ all three lengths are present and the ratio is φ^2 . $\square$

Direct computation confirms that $R(1/\varphi, N)$ takes exactly the two values φ and φ^2 for every N from 3 to 3000.

The metallic mean gap ratio theorem [7] places this result in its family. For the periodic multipliers $[0; a, a, a, \dots]$ the gap ratio eventually takes the $a + 1$ values $\mu_a - a + 1, \dots, \mu_a, \mu_a + 1$, where $\mu_a = (a + \sqrt{a^2 + 4})/2$ is the a -th metallic mean. The largest of these, $\mu_a + 1$, is $\varphi^2 \approx 2.618$ for $a = 1$, $2 + \sqrt{2} \approx 3.414$ for $a = 2$ and about 4.303 for $a = 3$. A noble multiplier eventually behaves like $1/\varphi$, with its gap ratio taking only values close to φ and φ^2 , and the multipliers with this two-valued behavior are exactly the noble ones [7]. A partial quotient of 2 that keeps recurring is enough to break it. Take a continued fraction of ones with a 2 at positions 4, 10, 19, 33, 55 and 90: over the first 30 000 key counts its gap ratio climbs as high as 3.62, and it exceeds φ^2 by more than 0.01 at 8275 of them.

For a hash table, R compares the widest empty stretch between neighboring keys on the circle with the narrowest. With $1/\varphi$, each new key falls into one of the largest slots and divides it in the golden ratio, and no slot is ever more than φ^2 times another.

4. Bucket Loads

The load of the fullest bucket is a different measurement. When each bucket holds many keys, it depends on how evenly the whole set of points fills every interval of length $1/m$, and the gap ratio alone does not decide it. We measured it directly.

For each of 169 prime table sizes m from 7 to 1021, we hashed $N = 1000m$ keys with each multiplier and recorded the fullest bucket's count divided by the average count of 1000. Primes avoid artifacts from common factors. Table I gives the results, and Figure 1 shows them across table sizes.

Table I. Fullest bucket divided by the average, over 169 prime table sizes, $N = 1000m$ keys.

Multiplier	Worst	P95	Mean	Best	Std
$1/\varphi$	1.0030	1.0030	1.0019	1.0010	0.0005
$1/(1 + \varphi) = 1/\varphi^2$	1.0030	1.0030	1.0019	1.0010	0.0005
$[0; 3, 1, 1, \dots] = 1/(2 + \varphi)$	1.0030	1.0030	1.0021	1.0010	0.0005
$[0; 5, 1, 1, \dots]$	1.0030	1.0030	1.0020	1.0010	0.0005
$\sqrt{2} - 1$	1.0030	1.0030	1.0021	1.0010	0.0005
$1/e$	1.0050	1.0050	1.0030	1.0010	0.0009
$1/\pi$	1.0710	1.0650	1.0300	1.0000	0.0195

The rows for $1/\varphi$ and $1/\varphi^2$ agree because $1/\varphi^2 = 1 - 1/\varphi$: the two place every key at mirror-image positions, so their fullest buckets are the same.

With 1000 keys per bucket the worst case moves in steps of 0.001, one key at a time, and at that resolution $1/\varphi$, $\sqrt{2} - 1$ and the noble multipliers all reach the same worst case, 1.0030, three keys above the average. The mean separates them only slightly: 1.0019 for $1/\varphi$ against 1.0021 for $\sqrt{2} - 1$. The clear separation is between bounded and unbounded partial quotients. The multiplier $1/e$, whose partial quotients $[0; 2, 1, 2, 1, 1, 4, 1, 1, 6, \dots]$ grow slowly, reaches 1.0050.

The multiplier $1/\pi$, with quotients as large as 292, reaches 1.0710, and its standard deviation across table sizes is 39 times that of $1/\varphi$.

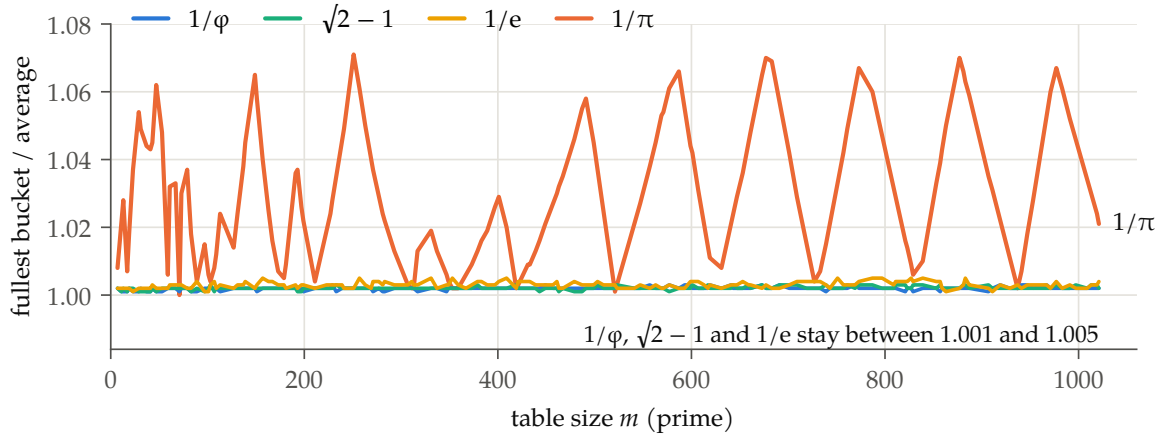


Figure 1. Fullest bucket divided by the average, for each prime table size m up to 1021, with $1000m$ keys. The multiplier $1/\pi$ swings between perfect balance and 7 percent overload as m changes; the three multipliers with bounded partial quotients stay between 1.001 and 1.005.

That variability is the practical failure. A programmer cannot choose m to dodge the bad table sizes of $1/\pi$ in advance. At $m = 71$, a divisor of 355 from the approximation $\pi \approx 355/113$, it balances its buckets perfectly, and at $m = 251$ it overloads the fullest bucket by 7.1 percent.

5. The Gap Envelope

Figure 2 follows the largest gap $G(\alpha, N)$ for N up to 20 000, scaled as the envelope $E(\alpha, N) = N \cdot G(\alpha, N)$.

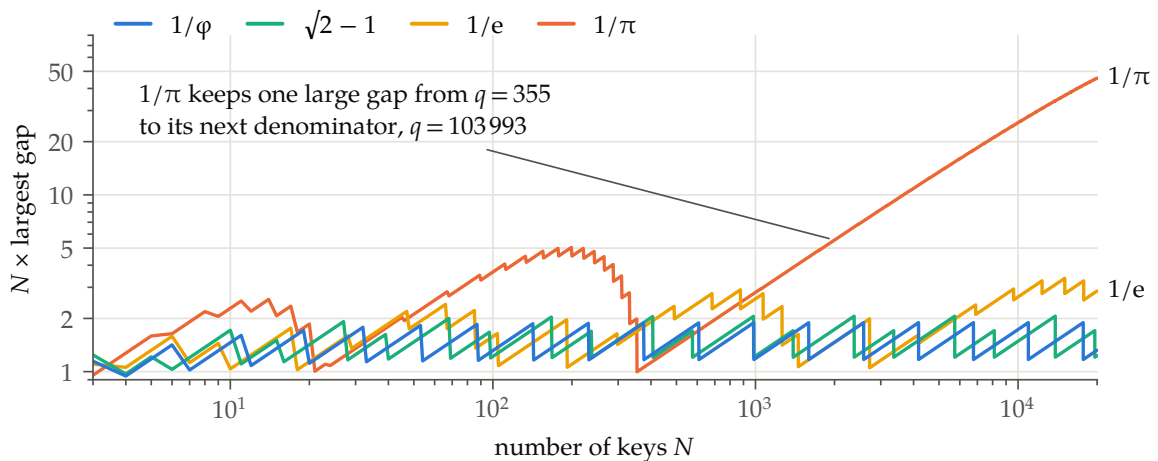


Figure 2. The gap envelope $N \cdot G(\alpha, N)$ on logarithmic axes. For $1/\varphi$ it stays between 0.94 and 1.89. For $1/\pi$ it climbs along each plateau and, after $N = 355$, keeps climbing until the next denominator, 103 993.

At the convergent denominators the envelope settles to a limit. For $1/\varphi$, $q_n G(1/\varphi, q_n) \rightarrow \varphi^2/\sqrt{5} \approx 1.1708$. For $\sqrt{2} - 1$ the limit is $(1 + \sqrt{2})/2 \approx 1.2071$. Just before each drop the

envelope of $1/\varphi$ peaks near $\varphi^3/\sqrt{5} \approx 1.894$.

This quantity does not single out the golden ratio: for the periodic multiplier $[0; 4, 4, 4, \dots]$ the limit is $(\varphi^3 + 1)/\sqrt{20}$, and since $\varphi^3 + 1 = 2\varphi^2$, that is again $\varphi^2/\sqrt{5}$. The gap ratio of Section 3 does.

For $1/\pi = [0; 3, 7, 15, 1, 292, \dots]$ the denominators are 1, 3, 22, 333, 355, 103 993, $\dots$. Each large partial quotient holds one large gap in place for a long run of keys, and the envelope climbs along it; after $N = 355$ it keeps climbing until $N = 103\,993$.

6. Noble Multipliers

Noble multipliers agree with $1/\varphi$ in the long run. We tested six of them, chosen so that no two are equal and no two are mirror images: $[0; 1, 1, 1, \dots] = 1/\varphi$, $[0; 3, 1, 1, \dots]$, $[0; 4, 1, 1, \dots]$, $[0; 5, 1, 1, \dots]$, $[0; 2, 2, 1, 1, \dots]$ and $[0; 3, 2, 1, 1, \dots]$. Over the 77 prime table sizes from 500 to 1021, with $1000m$ keys, their mean overloads of the fullest bucket lie between 2.06 and 2.21 keys above the average, with a standard deviation of 0.05 keys between the six. For comparison, $\sqrt{2} - 1$ averages 2.32, $1/e$ averages 3.05, and $1/\pi$ averages 35.8 (Figure 3).

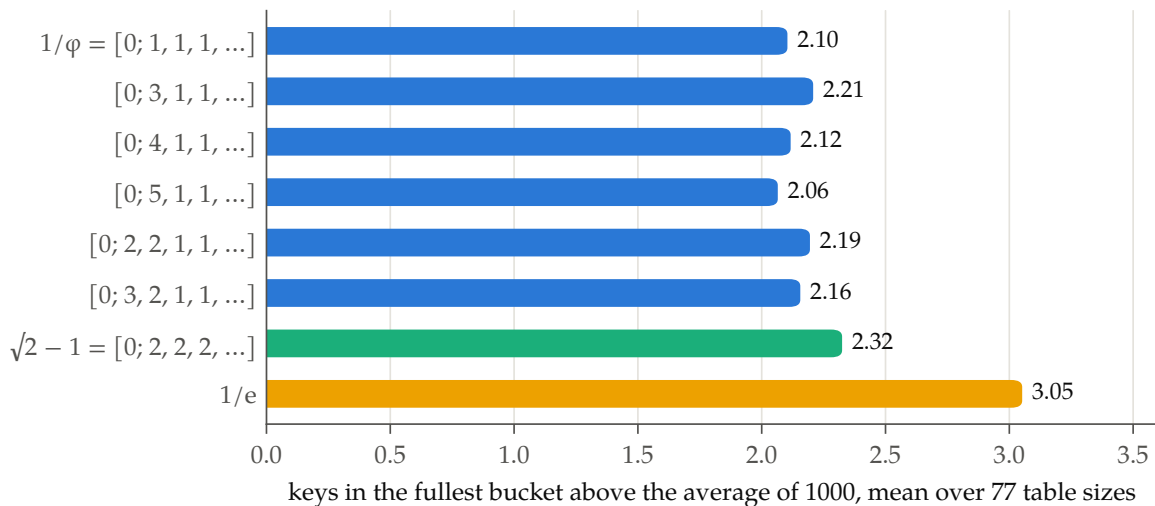


Figure 3. Mean number of keys in the fullest bucket above the average of 1000, over the 77 prime table sizes from 500 to 1021. The six noble multipliers form one cluster. The multiplier $1/\pi$, at 35.8, is off the scale.

The leading partial quotients change the details at small table sizes, and the common tail of ones decides the long run. The same pattern appears in the LZ78 phrase counts of Sturmian sequences [5].

7. Discussion

7.1 Variance

The measurements say more about variance than about the worst case. Multipliers with bounded partial quotients give a fullest bucket that hardly changes with the table size (standard deviation 0.0005), while $1/\pi$ moves between perfect balance and 7 percent overload

(0.0195).

7.2 The Diophantine Axis

Hash uniformity is a question of distribution, the same kind of question as compression resistance of Sturmian sequences [5] and quasi-Monte Carlo discrepancy [6]. The deciding property is the size of the partial quotients, which is Diophantine. The algebraic identity $\varphi^2 = \varphi + 1$ enters only through Theorem 1, where it makes every gap a power of $1/\varphi$; there is no frequency mixing in hashing.

7.3 Practical Implications

For practitioners the result supports Knuth’s recommendation, with a clear account of what it buys. The golden ratio multiplier $1/\varphi \approx 0.6180339887$ spaces keys more evenly than any other multiplier with a periodic continued fraction, only the other noble multipliers match it in the long run, and in the tests it and the other nobles have the lowest mean overload, about 0.2 percent with 1000 keys per bucket. Any multiplier with small partial quotients, such as $\sqrt{2} - 1$, reaches the same worst case in these tests, and all of them avoid the pathological table sizes that large partial quotients produce.

The analysis also gives a simple test for a proposed multiplier: compute its continued fraction and look for large partial quotients. Each large quotient predicts overloaded buckets at table sizes near the corresponding convergent denominator.

7.4 Relation to Existing Results

The ingredients are classical. The Three-Distance Theorem dates to 1958 [2], Hurwitz’s theorem to 1891 [3], the continued fraction description of the gaps is textbook material [4], and Knuth’s recommendation dates to 1973 [1]. What this paper adds is the exact two-valued gap ratio for $1/\varphi$, and measurements that separate what the golden multiplier guarantees from what it shares with other multipliers.

The results concern irrational multipliers applied to keys with unknown or adversarial structure. For finite key sets with known structure, tailored hash functions may do better than any multiplicative hash.

8. Conclusion

Knuth’s choice of $1/\varphi$ as the multiplier for multiplicative hashing has an exact justification: its keys cut the circle into gaps whose largest is always φ or φ^2 times the smallest, the tightest spread of any multiplier with a periodic continued fraction and one that only the noble multipliers match in the long run. Bucket loads follow the same order in the mean. In the worst case, other multipliers with small partial quotients do as well at 1000 keys per bucket, and multipliers with large partial quotients do far worse.

Changes from the first edition

1. Theorem 1. The first edition stated that $1/\varphi$ minimizes the worst-case bucket load over all table sizes. Its own Table I shows $\sqrt{2} - 1$ and the noble multipliers reaching the same worst case, 1.0030. The theorem is replaced by the gap-ratio result, which is proved, and the load results are reported as measurements.
2. Lemmas 2 and 3. The largest gap at $N = q_n$ is $\|q_{n-1}\alpha\| + \|q_n\alpha\|$, not $1/q_{n-1}$, and for $1/\varphi$ the product $q_n G(\alpha, q_n)$ tends to $\varphi^2/\sqrt{5} \approx 1.1708$, not to φ . The multiplier $[0; 4, 4, 4, \dots]$ reaches the same limit, so this quantity cannot single out $1/\varphi$.
3. Experiment 2. The first edition said the product at the convergents converges to about 1.1708 and also that it approaches φ . The limit is $1.1708 = \varphi^2/\sqrt{5}$.
4. Experiment 3. The six noble numbers listed were four: $1/(1 + \varphi)$ equals $1/\varphi^2$, and $[0; 3, 1, 1, \dots]$ equals $1/(2 + \varphi)$. Also, $1/\varphi^2$ gives the mirror image of $1/\varphi$. The experiment was rerun with six different nobles, no two of them mirror images. The first edition's Table II values could not be reproduced with 1000 keys per bucket; the rerun states its setup.
5. Figures 1 to 3 were missing from the first edition's file and are now included.
6. Table sizes. There are 169 primes between 7 and 1021, not 170.
7. Section 5.1. The first edition placed the perfect balance of $1/\pi$ near $m = 355$. Among the prime table sizes it occurs at $m = 71$, a divisor of 355; the primes next to 355 are not balanced.
8. Practical implications. The first edition called the advantage of $1/\varphi$ over $\sqrt{2} - 1$ structurally guaranteed. Every multiplier with bounded partial quotients avoids pathological table sizes, so that claim is withdrawn.

Acknowledgments

The mathematical analysis, computational verification, and portions of the manuscript text were developed with the assistance of Claude (Anthropic), a large language model. The corrections in this edition were found and checked by computation with Claude (Anthropic). The author is responsible for the research direction, theoretical framing, and all claims made in this paper.

References

- [1] D. E. Knuth, *The Art of Computer Programming, Vol. 3: Sorting and Searching*, 2nd ed. Addison-Wesley, 1998.
- [2] V. T. Sós, "On the distribution mod 1 of the sequence $n\alpha$," *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, vol. 1, pp. 127–134, 1958.
- [3] A. Hurwitz, "Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche," *Math. Ann.*, vol. 39, pp. 279–284, 1891.
- [4] A. Ya. Khinchin, *Continued Fractions*, 3rd ed. Chicago, IL: University of Chicago Press, 1964.

- [5] J. Alfredo, "Maximal LZ78 complexity among Sturmian sequences: noble number extremality and epoch packing," Scalar Resonance Research Program, 2026.
- [6] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia, PA: SIAM, 1992.
- [7] J. Alfredo, "The metallic mean gap ratio theorem: a characterization of noble numbers via the Three-Distance Theorem," Scalar Resonance Research Program, 2026.