On Finding Best Rational Approximations

In systems utilizing inexpensive microcontrollers, a common technique for efficiently implementing a linear scaling $f(x) = rx$ from an integer domain to an integer range is to approximate a real number r by a rational number h/k. The linear scaling can then be implemented using a single integer multiplication instruction followed by a single integer division instruction, i.e.[1]

$$g(x) = \left\lfloor \frac{hx}{k} \right\rfloor . \tag{1}$$

As an example of this technique, using an inexpensive microcontroller that has machine instructions to multiply by a single unsigned byte or divide by a single unsigned byte, we might wish to convert from integral miles-per-hour to integral kilometers-per-hour. We would want a rational number h/k which is as close as possible to the ideal conversion factor (1.6093), but with the restriction that neither the numerator nor the denominator can be larger than 255 (the maximum single-byte value accomodated by the machine instructions). The best rational approximation h/k under these constraints is 243/151—but how do we find this rational number? How do we go from the ideal conversion factor (1.6093) to the best rational number (243/151)?

In this article, we consider the general problem of choosing a rational number h/k as close as possible to an arbitrary real number r, subject to the constraints $h \leq HMAX$ and $k \leq KMAX$. We require constraints on the numerator h and denominator k because machine instructions are always limited in the size of the operands they can accomodate.

The problem of economically choosing a rational number as close as possible to an arbitrary real number is a counterintuitively difficult problem. The arrangement of rational numbers on the real number line is a topic from number theory (a branch of mathematics). In this article, we present a concrete procedure for finding best rational approximations.

By *rational number*, we mean a number that can be expressed as a fraction, such as 243/151. By *irreducible* rational number, we mean a rational number that is in lowest terms—no common factors can be divided out of the numerator and denominator. By *best rational approximations*, we mean the two irreducible

[1] Throughout this article, we use the notation $\lfloor\ \rfloor$ to denote the *floor* function, where $\lfloor z \rfloor$ is the largest integer not larger than z. For example, $\lfloor 1.9 \rfloor = 1$ and $\lfloor 2 \rfloor = 2$. In equation (1), we use the floor function to indicate that the remainder of the division is discarded.

rational numbers of the form h/k enclosing r (the number we wish to approximate) which meet the constraints $h \leq HMAX$ and $k \leq KMAX$.

The Farey Series

Although this article considers rational numbers of the form h/k with both numerator and denominator constrained ($h \leq HMAX$ and $k \leq KMAX$), it is helpful to first consider what rational numbers we can form with only a constrained denominator ($k \leq KMAX$). The ordered set of irreducible rational numbers in the interval [0,1][2] that can be formed when only the denominator is constrained ($k \leq KMAX$) is a standard set in number theory, and is called *The Farey Series*. The *order* of the Farey series is the maximum denominator allowed. The Farey series of order N is denoted F_N. For example, the Farey series of order 5, denoted F_5, is

$$F_5 = \left\{ \frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1} \right\}. \tag{2}$$

Given two consecutive terms in F_N (h_{j-2}/k_{j-2} and h_{j-1}/k_{j-1}), the next term can be calculated using the following equations ([9, p. 83]):

$$h_j = \left\lfloor \frac{k_{j-2} + N}{k_{j-1}} \right\rfloor h_{j-1} - h_{j-2} \tag{3}$$

$$k_j = \left\lfloor \frac{k_{j-2} + N}{k_{j-1}} \right\rfloor k_{j-1} - k_{j-2} \tag{4}$$

Similarly, given two consecutive terms in F_N (h_{j+1}/k_{j+1} and h_{j+2}/k_{j+2}), the previous term can be calculated using:

$$h_j = \left\lfloor \frac{k_{j+2} + N}{k_{j+1}} \right\rfloor h_{j+1} - h_{j+2} \tag{5}$$

$$k_j = \left\lfloor \frac{k_{j+2} + N}{k_{j+1}} \right\rfloor k_{j+1} - k_{j+2} \tag{6}$$

[2] The Farey series is defined to be in the interval [0,1] for theoretical reasons. It is easy to see that the Farey series "repeats"—each rational number h/k in [0,1] has a counterpart $(ik + h)/k$ in any interval bounded by two consecutive integers, $[i, i + 1]$. Thus, information about the distribution of the Farey rational numbers in [0,1] gives complete information about the distribution everywhere, which is why mathematicians are only interested in the interval [0,1]. In this article, we do abuse the proper nomenclature somewhat and refer to terms outside the interval [0,1] as Farey terms. We'd like to assure readers that this abuse is not harmful, and that all of the results presented in this article apply for all non-negative rational numbers.

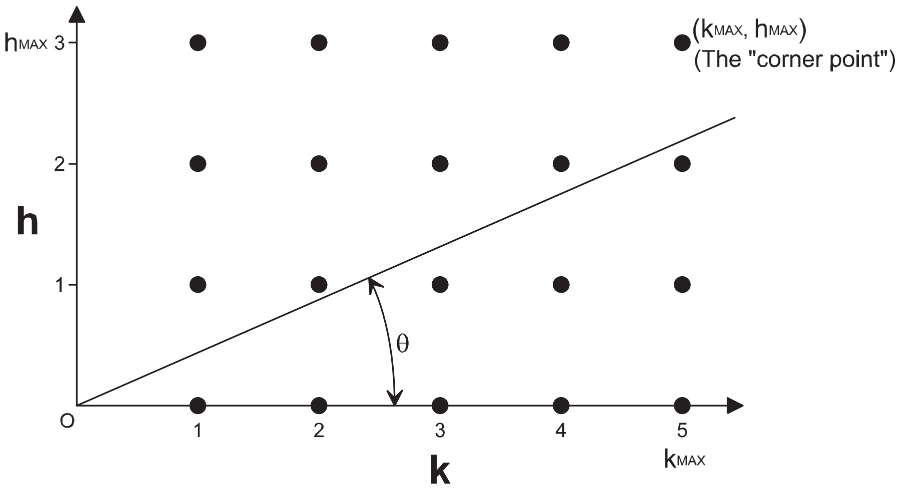

Figure 1: Graphical Interpretation Of Rational Numbers h/k That Can Be Formed With $h \leq HMAX$ And $k \leq KMAX$

Note that the derivation of equations (3) through (6) absolutely requires that the rational numbers used as input be irreducible, and absolutely guarantees that the result produced will be irreducible.

Using the above equations, we can build the ordered set of irreducible rational numbers with a certain maximum denominator, starting at any integer. At an integer i in the Farey series of order N, $i/1$ and $(iN+1)/N$ are always two consecutive Farey terms. Thus, equations (3) through (6) can be used to build as many terms as needed in either direction, starting at any convenient integer.

Rectangular Regions

The previous section showed that if only the denominator is constrained, $k \leq KMAX$, the ordered set of irreducible rational numbers that can be formed is the Farey series of order $KMAX$. However, it isn't intuitively obvious what set of rational numbers can be formed when *both* the numerator and denominator are constrained.

The complete set of rational numbers that can be formed subject to the constraints $h \leq HMAX$ and $k \leq KMAX$ has a convenient and intuitive graphical interpretation (Fig. 1). In Fig. 1, each dot corresponds to a rational number—not necessarily irreducible—that can be formed with $h \leq 3$ and $k \leq 5$.

With a little thought, the following properties are apparent:

- The value of the rational number h/k is monotonic with respect to the angle of the ray drawn from the origin to the corresponding dot. $h/k = tan\ \theta$, and $\theta = tan^{-1} h/k$.

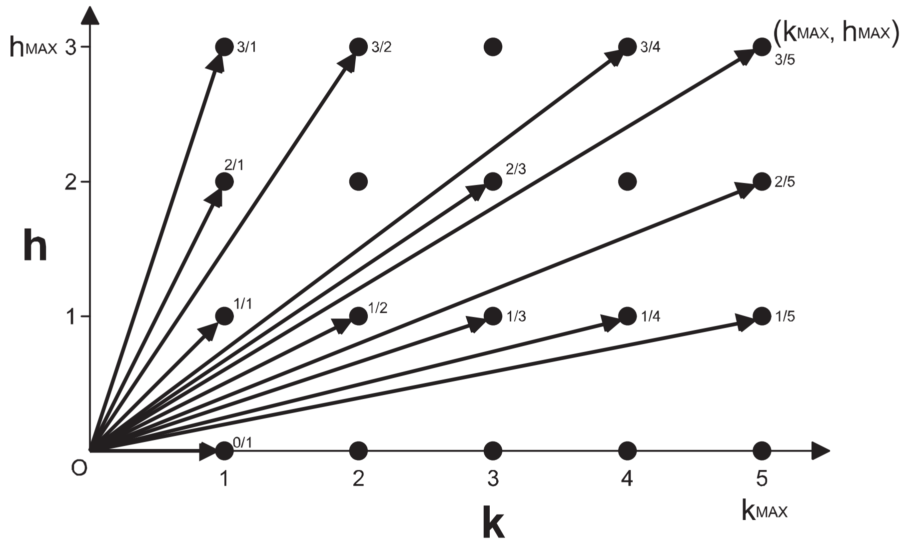

Figure 2: Graphical Method Of Constructing Set Of Irreducible Rational Numbers h/k That Can Be Formed With $h \leq HMAX$ And $k \leq KMAX$

- Only irreducible rational numbers have dots that are directly visible from the origin (without an intervening dot).
- The complete ordered set of irreducible rational numbers that can be formed can be graphically constructed by sweeping a ray from the origin starting at $\theta = 0$ through the angles $0 \leq \theta < \pi/2$ and recording the coordinates of each dot directly visible from the origin.

Fig. 2 illustrates the process of graphical construction. In the figure, a ray is drawn from the origin to each dot that can be reached without an intervening dot. The dots are labeled with the irreducible rational numbers they represent.

The strength of Figs. 1 and 2 is that they show clearly how to deal with the constraint on the numerator, $h \leq HMAX$. As we sweep a ray from the origin through the angles $0 \leq \theta \leq tan^{-1}3/5$ (i.e., up to the corner point), it is clear that the dots directly visible from the origin are simply the terms of F_5. However, for the angles $tan^{-1}3/5 < \theta < \pi/2$ (i.e., after the corner point), it is clear that a different series is involved, since many of the terms of F_5 are not available.

From symmetry in Figs. 1 and 2, it is clear that after the corner point, we are encountering the terms of F_3, but inverted (i.e., their reciprocals) and in the reverse order. (Imagine exchanging axes in the figures.) Thus, it is clear that the ordered set of irreducible rational numbers that can be formed subject to the constraints $h \leq 3$ and $k \leq 5$ is a portion of F_5 concatenated with the reciprocals of the reverse-ordered terms in a portion of F_3. This observation suggests an algorithm for generating the complete ordered set of irreducible rational numbers that can be formed subject to the constraints $h \leq HMAX$ and $k \leq KMAX$, which we present as Algorithm 1.

Algorithm 1—*Algorithm For Constructing Ordered Set Of Irreducible Rational Numbers In A Rectangular Region Of The Integer Lattice*

- Build F_{KMAX} from 0/1 through $HMAX/KMAX$ (or its reduced equivalent), using equations (3) and (4). (In Figs. 1 and 2, these are the dots up through the corner point.)
- Build F_{HMAX} from 1/HMAX through $KMAX/HMAX$ (or its reduced equivalent), using equations (3) and (4). Take the reciprocal of each term, and reverse the order of the terms. (In Figs. 1 and 2, these are the dots after the corner point.)
- Concatenate the results from the two steps above.

For example, to build the ordered set of irreducible rational numbers shown in Figs. 1 and 2 ($HMAX = 3$, $KMAX = 5$) using Algorithm 1, first build F_5 from 0/1 through HMAX/KMAX:

$$F_5 = \left\{ \frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \ldots \right\}. \tag{7}$$

Then build F_3 from 1/HMAX to KMAX/HMAX:

$$F_3 = \left\{ \ldots, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, \ldots \right\}, \tag{8}$$

and invert and reverse the order of the terms in F_3 to yield

$$F_{\overline{3}} = \left\{ \ldots, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{3}{1}, \ldots \right\}. \tag{9}$$

Finally, concatenate equations (7) and (9) to yield equation (10), the complete set of irreducible rational numbers that can be formed subject to the constraints $h \leq 3$ and $k \leq 5$.

$$F_{5,\overline{3}} = \left\{ \frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{3}{1} \right\}. \tag{10}$$

Note most importantly that Algorithm 1 reduces a problem with two constraints ($h \leq HMAX$ and $k \leq KMAX$) to a problem with one constraint, solved twice. Thus equations (3) through (6) can be applied to easily generate all of the irreducible rational numbers that can be formed under the constraints

$h \leq HMAX$ and $k \leq KMAX$. Since Algorithm 1 provides a way to generate, in order, all of the irreducible rational numbers with a certain maximum numerator and denominator, it also provides a way to find the two rational numbers that enclose the number (r) we are trying to approximate. We can simply generate rational numbers, in order, until r is enclosed.

It is also noteworthy that F_3 and F_5 (the examples used so far) are trivial, and most readers can construct these Farey series in their head. Naturally, the real power of equations (3) through (6) is that these relationships give a closed-form way to construct higher-order Farey series, such as F_{255}, $F_{65,6535}$, and $F_{4,294,967,295}$—the most typical Farey series encountered in embedded microcontroller software.

The Continued Fraction Algorithm

When Farey series of small order ($HMAX$ and $KMAX$ of several hundred or less) are involved, the techniques presented so far are adequate for finding best rational approximations. It is very easy and practical to construct a simple computer program to generate in order all irreducible rational numbers that meet the constraints and then choose the two that enclose r. This is a practical technique because the number of rational numbers that need to be constructed under the constraints will be several hundred thousand at most.

However, when Farey series of higher order are involved, the techniques presented so far are not adequate. A result from number theory is that a Farey series of order N contains about $3N^2/\pi^2$ terms. With powerful microcontrollers that have machine instructions to multiply and divide very large integers, it may be necessary to find best rational approximations in the Farey series of up to order $2^{32} - 1$. A simple calculation would show that in the Farey series of order $2^{32} - 1$, enumeration of all the rational numbers between two consecutive integers would require at least tens of thousands of years, even for the fastest personal computer. We need to develop a technique that doesn't involve sifting through all available choices.[3]

An elegant $O(log\ N)$ solution is available from the study of *continued fractions*, and we present this solution as Algorithm 2. Because the solution is $O(log\ N)$, it can be applied to find best rational approximations even in Farey series of very large order, and it can easily be applied by hand if necessary.

A finite simple continued fraction is a fraction of the form

[3] To be more precise about it ... constructing (or linearly searching) F_N between two consecutive integers is $O(N^2)$. There are several refinements that will lead to an $O(N)$ algorithm, but even an $O(N)$ algorithm isn't acceptable for the Farey series of order $2^{32} - 1$. The continued fraction algorithm that we present as Algorithm 2 is $O(log\ N)$.

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ldots + \cfrac{1}{a_n}}}} \tag{11}$$

where a_0 is a non-negative integer and each of the other coefficients $a_1, a_2, \ldots, a_n$ is a positive integer. Algorithm 2 works by obtaining the continued fraction representation of the number r_I to which we'd like to obtain a best rational approximation.

With very little further ado or explanation, we present Algorithm 2. Note in the algorithm that the integers a_k calculated are the coefficients in equation (11) of the continued fraction representation of a/b. Note also that the rational numbers p_k/q_k calculated are called the *convergents*[4] of the continued fraction.

Algorithm 2—*Continued Fraction Algorithm For Finding Best Rational Approximations*

- Express the number to which best rational approximations are needed, r, as a rational number a/b with a high degree of precision (far more precision than is needed for the approximation). This rational number a/b does not have to be reduced. The most practical technique in most cases is to choose a to be the digits from the display of a pocket calculator (without the decimal point), and b to be a power of ten to place the decimal point correctly. For example, to approximate π, choose a/b = 314,159,265,359 / 100,000,000,000.

- Determine whether $r > HMAX/KMAX$. (If $r > HMAX/KMAX$, it is beyond the "corner point" in the sense suggested by Fig. 1.) If $r < HMAX/KMAX$, proceed with $MAX = KMAX$ and $a/b = r$. However, if $r > HMAX/KMAX$, we must proceed with $MAX = HMAX$ and $a/b = 1/r$. We will then need to use the reciprocals of the results produced by the rest of the algorithm.

- Make the following initial variable assignments:

  $k := -1$
  $divisor_{-1} = a$
  $remainder_{-1} = b$

[4]Convergents have many special properties not mentioned in this article. For example, the convergents p_k/q_k alternate on both sides of a/b and get ever-closer. It is also a property that for any convergent p_k/q_k, there is no rational number with a smaller denominator that is closer to a/b. We should also mention that all convergents are provably irreducible, and that the result obtained from equation (12) will also be provably irreducible. The final convergent (if Algorithm 2 is carried out until a zero remainder) will be a/b in lowest terms.

$p_{-1} = 1$
$q_{-1} = 0$

- REPEAT
  $k := k + 1$
  $dividend_k := divisor_{k-1}$
  $divisor_k := remainder_{k-1}$
  $a_k := dividend_k \ DIV \ divisor_k$
  $remainder_k := dividend_k \ MOD \ divisor_k$
  IF $(k = 0)$ $p_0 := a_0 = \lfloor r_I \rfloor$ ELSE $p_k := a_k p_{k-1} + p_{k-2}$
  IF $(k = 0)$ $q_0 := 1$ ELSE $q_k := a_k q_{k-1} + q_{k-2}$
  UNTIL $(remainder_k = 0$ OR $q_k > MAX)$

- If the loop terminates without finding $q_k > MAX$, the rational number a/b was already in the set of rational numbers that could be formed under the constraints (but perhaps a/b was not reduced and so this was not obvious). In this case, use the final p_k/q_k as a best rational approximation, as this is the reduced form of a/b.

- If the loop terminates with $q_k > MAX$, the number a/b was *not* already in the set of rational numbers that could be formed under the constraints. In this case, p_{k-1}/q_{k-1} is one best approximation under the constraints to a/b, and the other best approximation is

$$\frac{\left\lfloor \dfrac{N - q_{k-2}}{q_{k-1}} \right\rfloor p_{k-1} + p_{k-2}}{\left\lfloor \dfrac{N - q_{k-2}}{q_{k-1}} \right\rfloor q_{k-1} + q_{k-2}}. \tag{12}$$

A Design Example

We present one design example to illustrate the method. Assume we wish to rationally approximate $1/\pi$ and π on a microcontroller that has machine instructions to multiply by integers as large as 10,000, and divide by integers as large as 20,000 (thus $HMAX$ = 10,000 and $KMAX$ = 20,000). For the example, we will find the best rational approximations to $1/\pi$ and π under these constraints.

The first step in Algorithm 2 is to express the number we'd like to approximate as a rational number of high precision. A typical scientific calculator supplies 3.14159265359 as π, so we'll use the rational number a/b = 314,159,265,359 / 100,000,000,000 as the value of π and a/b = 100,000,000,000 / 314,159,265,359 as the value of $1/\pi$.

We note that $1/\pi < HMAX/KMAX$ (i.e., below the corner point), so all of the terms of $F_{20,000}$ are available to choose as rational approximations. Applying the numerical steps in Algorithm 2 yields Table 1.

Table 1: Application Of Algorithm 2, Using a/b = 100,000,000,000 / 314,159,265,359 As The Value Of $1/\pi$

Index (k)	$dividend_k$	$divisor_k$	a_k	$remainder_k$	p_k	q_k
-1	N/A	100,000,000,000	N/A	314,159,265,359	1	0
0	100,000,000,000	314,159,265,359	0	100,000,000,000	0	1
1	314,159,265,359	100,000,000,000	3	14,159,265,359	1	3
2	100,000,000,000	14,159,265,359	7	885,142,487	7	22
3	14,159,265,359	885,142,487	15	882,128,054	106	333
4	885,142,487	882,128,054	1	3,014,433	113	355
5	882,128,054	3,014,433	292	1,913,681	33,102	103,993

Note in Table 1 that the algorithm is applied only until q_k exceeds the maximum denominator $MAX = KMAX = 20,000$. Note also that at termination, $k = 5$.

By Algorithm 2, $p_{k-1}/q_{k-1} = 113/355$ is one best approximation to $1/\pi$ under the constraints. Applying equation (12) (with $k = 5$ and $MAX = 20,000$) to obtain the other best approximation yields 6,321/19,858.

Obtaining the best approximations to π subject to $h \leq 10,000$ and $k \leq 20,000$ is a less intuitive process. Note that $\pi > HMAX/KMAX$ (i.e., above the corner point), so we must find the two best approximations to $1/\pi$ in the Farey series of order $HMAX = 10,000$, and then take the reciprocals of these two best approximations.

To find the best approximations to $1/\pi$ in $F_{10,000}$, Table 1 can be reused. As before, Algorithm 2 would terminate with $k = 5$, because $q_5 = 103,993$ is the first value of q with a denominator greater than $10,000$. As before $p_{k-1}/q_{k-1} = 113/355$ is one best approximation to $1/\pi$. Applying equation (12) (with $k = 5$ and $N = 10,000$) to obtain the other best approximation yields 3,157/9,918. Calculating the reciprocals of these two best approximations to $1/\pi$ yields 355/113 and 9,918/3,157 as the two best approximations to π, subject to the constraints $h \leq 10,000$ and $k \leq 20,000$.

Software Solutions

Software and Excel spreadsheets to perform the calculations described in this article are available for free download from the Internet at *www.embedded.com/code.html*. The spreadsheets `FAREY_FORWARD.XLS` and `FAREY_BACKWARD.XLS` build the Farey series forward and backward from an integer, using equations (3) through (6). The spreadsheet `CF_METHOD.XLS` implements the continued fraction algorithm presented. These spreadsheets are believed to correctly operate in Farey series up to order at least 2^{24}.

More flexible and powerful software named *The Iju Tool Set* is distributed at the URL `ijutools.sourceforge.net`. This tool set includes a statically linked version of *Wish* (the classic Tcl/Tk interpreter) called *IjuConsole*, with arbitrary-size integer extensions built in. *IjuConsole* will operate in the Farey series of up to at least order 10^{1000} (but probably much higher). For example, entering the command

```
arbint brap [arbint const pi 1000] 1e20 1e20
```

will result in the output

```
2646693125139304345/842468587426513207,
```

which is the best rational approximation to π with numerator and denominator not exceeding 10^{20}. All of the available commands are documented in the book accompanying the tool set.

References

[1] Hardy, G.H. and Wright, E.M., *An Introduction To The Theory Of Numbers, Fifth Edition*, Oxford University Press, 1979.

[2] Harman, G., *Metric Number Theory*, Oxford University Press, 1998.

[3] Kargaev, P. and Zhigljavsky, A., *Approximation Of Real Numbers By Rationals: Some Metric Theorems*, Journal of Number Theory, 61, 209-225, 1966.

[4] Kargaev, P. and Zhigljavsky, A., *Asymptotic Distribution Of The Distance Function To The Farey Points*, Journal of Number Theory, 65, 130-149, 1967.

[5] Khinchin, A. Ya., *Continued Fractions*, University Of Chicago Press, 1964.

[6] LeVeque, William J., *Fundamentals Of Number Theory*, Dover Publications, 1977.

[7] Olds, C.D., *Continued Fractions*, The Mathematical Association Of America, 1963.

[8] Ore, Oystein, *Number Theory And Its History*, Dover Publications, 1948.

[9] Schroeder, M.R., *Number Theory In Science And Communication, Third Edition*, Springer, 1997.

[10] *Farey Series*,
`www.cut-the-knot.com/blue/Farey.html`

[11] *'C' Program For Best Rational Approximations To Real Numbers*,
`www.ics.uci.edu/~eppstein/numth/frap.c`

[12] *Rational Approximation Newsgroup Post*,
`www.math.niu.edu/~rusin/known-math/95/rationalize`

Biography And Contact Information

David T. Ashley, Joseph P. DeVoe, Karl Perttunen, and Cory Pratt are embedded system software engineers with Visteon in Dearborn, Michigan and 3Com in Chicago, Illinois. Anatoly Zhigljavsky is a number theorist at the University of Cardiff in Cardiff, England. The unusual collaboration between engineers and a mathematician leading to this article began in 2000 with two of the engineers discussing the best way to linearly scale bytes, and quickly realizing that nothing in an undergraduate engineering curriculum had prepared them to know exactly how rational numbers are arranged among the real numbers.

The authors may be contacted at the e-mail addresses below.

David T. Ashley: `dtashley@wayne.edu`
Joseph P. DeVoe: `jdevoe@visteon.com`
Karl Perttunen: `kperttun@visteon.com`
Cory Pratt: `cory_pratt@3com.com`
Anatoly Zhigljavsky: `zhigljavskyaa@cardiff.ac.uk`

Acknowledgements

We would like to gratefully acknowledge the assistance of Greg Bachelis, Robert Berman, Feng Lin, Nick Sahinidis, Adam Van Tuyl, Carl Schweiger, Ken Tindell, Steve Vestal, Bob Whitinger, David B. Stewart, Johan Bengtsson, Michael J. Burke, Mark Endicott, David Eppstein, David M. Einstein, Mircea Munteanu, Adam Gibson, Virgil, Cliff Stallings, Robert Kakos, Paulette Groen, Paula Smith, Bob Crosby, Klaus-Peter Zauner, Andrea Blome, Marilyn A. Ashley, Una Smith, Karsten Tinnefeld, Axel Franke, and Alexander Bogomolny in researching and preparing this article.

Do Not Include The Version Control Information Below In The Article.

The version control information below (generated automatically by CVS) is used only to assist during the editing cycle as the article is prepared for submission. Do not include this information in the article.

```
LaTeX compile date:  February 7, 2002.

CVS Information:
$Header: /smb/cvs_root/dtaipubs/dtaipubs/esp0101/esp0101a.tex,v 1.10 2001/12/28 17:03:16 dtashley Exp $
$Revision: 1.10 $
$Date: 2001/12/28 17:03:16 $
$Author: dtashley $
```