

Determination of the Best Constant in an Inequality of Hardy, Littlewood, and Pólya

T. C. Peachey and C. M. Enticott

CONTENTS

1. Introduction
 2. Determining the Best Constant in an Inequality
 3. Computation of the Functionals
 4. Numerical Evaluation
 5. Testing and Error Handling
 6. The Optimizations
 7. The Nimrod Tools
 8. Results and Discussion
- Acknowledgments
References

In 1934 Hardy, Littlewood, and Pólya generalized Hilbert's inequality to the case in which the parameters are not conjugate. Determination of the best constant in this generalization is still an unsolved problem. An experimental approach is presented that yields numerical values that agree with theory in the cases in which an exact answer is known. The results may be a guide to a further theoretical approach.

1. INTRODUCTION

Hilbert's integral inequality is the following well-known theorem.

Theorem 1.1. *If $p > 1$, $q = p/(p - 1)$, $f \in L_p[0, \infty]$, $g \in L_q[0, \infty]$, then*

$$\int_0^\infty \int_0^\infty \frac{f(u)g(v)}{u+v} du dv < B\left(\frac{1}{p}, \frac{1}{q}\right) \|f\|_p \|g\|_q \quad (1-1)$$

unless f or g is null.

Here the beta function $B(1/p, 1/q)$ is known to be the best possible "constant" in the sense that for any smaller constant, f and g can be found that will violate the inequality. A full discussion of this theorem is found in [Hardy et al. 52, Chapter 7]. In the same work, Theorem 340 generalized the inequality to the case in which p and q are not necessarily conjugate.

Theorem 1.2. *Suppose $p > 1$, $q > 1$, $p' = p/(p - 1)$, $q' = q/(q - 1)$, $0 < \lambda = 1/p' + 1/q' \leq 1$, $f \in L_p[0, \infty]$, and $g \in L_q[0, \infty]$. Then unless f or g is null,*

$$\int_0^\infty \int_0^\infty \frac{f(u)g(v)}{(u+v)^\lambda} du dv < C \|f\|_p \|g\|_q, \quad (1-2)$$

where C depends on p and q only.

2000 AMS Subject Classification: 26D15

Keywords: Integral inequality, distributed optimization

The authors noted that “the best value [of the constant] has not been found in the general case, and the problem of determining it appears to be difficult.”

The bound $C \leq K(p, q)$, where

$$K(p, q) = B^\lambda \left(\frac{1}{\lambda p'}, \frac{1}{\lambda q'} \right) = \left[\pi \operatorname{cosec} \left(\frac{\pi}{\lambda p'} \right) \right]^\lambda, \quad (1-3)$$

was first proved by Levin [Levin 36], and later more directly by Bonsall [Bonsall 51]. When p and q are conjugate, $\lambda = 1$ and Theorem 1.2 reduces to Theorem 1.1. So Levin’s bound (1-3) is known to give the best possible constant in this case. To this day, no smaller constant has been found, but no proof has been published showing that a smaller constant is not possible. (There has been some confusion on this matter. Finch [Finch 03] gives a thorough account.) See also [Peachey 03] for an alternative bound.

Henceforth we write $M(p, q)$ for the least possible value of C . In 1973 Walker [Walker 73] showed that as $p \rightarrow 1$ for fixed q , $M(p, q)/K(p, q) \rightarrow 1$. Symmetrically the same conclusion holds if $q \rightarrow 1$ for fixed p .

In terms of parameters $1/p'$ and $1/q'$ the allowed parameters are given by the triangular region

$$0 < \frac{1}{p'}, \quad 0 < \frac{1}{q'}, \quad \frac{1}{p'} + \frac{1}{q'} \leq 1. \quad (1-4)$$

The boundaries $1/p' = 0$ and $1/q' = 0$ correspond to $p \rightarrow 1$ and $q \rightarrow 1$, the cases considered by Walker. The other boundary corresponds to the case $\lambda = 1$, which is the original Hilbert inequality. So we know that $M(p, q) \rightarrow K(p, q)$ near all three boundaries of this region. Hence it is a reasonable hypothesis that Levin’s constant $K(p, q)$ is the best possible throughout.

The left side of (1-2) may be rewritten as

$$\int_0^\infty g(v) dv \int_0^\infty \frac{f(u)}{(u+v)^\lambda} du = \int_0^\infty g(v)F(v) dv, \quad (1-5)$$

say, where $g \in L_q$ and $F \in L_{q'}$. So (1-2) is $\int_0^\infty g(v)F(v) dv < C\|f\|_p\|g\|_q$ for all $g \in L_q$. Using the converse of Hölder’s inequality, [Hardy et al. 52, Proposition 191], shows that $\|F\|_{q'} < C\|f\|_p$. Conversely, the direct Hölder inequality shows that $\|F\|_{q'} < C\|f\|_p$ implies (1-2). Hence (replacing q' by r) Theorem 1.2 is equivalent to the following.

Theorem 1.3. *Suppose $p > 1$, $p' = p/(p - 1)$, $1/p' < \lambda \leq 1$, $1/r = \lambda - 1/p'$, and $f \in L_p$. Then for*

$$F(w) = \int_0^\infty \frac{f(u)}{(u+w)^\lambda} du, \quad (1-6)$$

one has

$$\|F\|_r < C\|f\|_p \quad (1-7)$$

unless f is null. Here C depends on p and λ only and its least value will be the same as that for C in (1-2).

This is a more convenient formulation for our purposes. It shows that the best constant will be the operator norm of the generalized Stieltjes transform, (1-6), when it maps from L_p to L_r .

This paper reports lower bounds on the constant C for a variety of values of p and λ computed using numerical evaluation on a cluster of computers.

2. DETERMINING THE BEST CONSTANT IN AN INEQUALITY

The task of proving that a given C in (1-2) is the least possible constant requires an experimental approach; given $\epsilon > 0$, finding “test functions” f_η and g_η such that

$$\int_0^\infty \int_0^\infty \frac{f_\eta(u)g_\eta(v)}{(u+v)^\lambda} du dv > (C - \epsilon)\|f_\eta\|_p\|g_\eta\|_q. \quad (2-1)$$

Note that this cannot be achieved for all ϵ by single functions f_η and g_η since then

$$\int_0^\infty \int_0^\infty \frac{f_\eta(u)g_\eta(v)}{(u+v)^\lambda} du dv \geq C\|f_\eta\|_p\|g_\eta\|_q,$$

which violates the strictness of the inequality. Instead infinite families of functions are required, containing members that satisfy (2-1) for each $\epsilon > 0$.

The case $\lambda = 1$ was settled, [Hardy et al. 52, §9.5], using the families

$$f_\eta(u) = \begin{cases} u^{-(1+\eta)/p} & \text{for } u \geq 1, \\ 0 & \text{elsewhere,} \end{cases}$$

$$g_\eta(v) = \begin{cases} v^{-(1+\eta)/q} & \text{for } v \geq 1, \\ 0 & \text{elsewhere,} \end{cases}$$

where η is positive. Walker’s proof of the case $p \rightarrow 1$ used

$$f_\eta(u) = \begin{cases} u^{-(1+\eta)/p} & \text{for } u \geq 1, \\ 0 & \text{elsewhere,} \end{cases}$$

$$g_\eta(v) = \begin{cases} v^{-(1+1/\eta)/q} & \text{for } v \geq 1, \\ 0 & \text{elsewhere,} \end{cases}$$

again with $\eta > 0$. In both cases the argument required $\eta \rightarrow 0$ as $\epsilon \rightarrow 0$, so the integral defining $\|f_\eta\|_p$ approached

a divergent integral. Generally one expects limiting extremes to settle a question of the best constant. See for example [Peachey et al. 99].

Since we consider Theorem 1.3, only a single family of test functions will be required. In this paper we report results using the family

$$f_{\mu,\nu}(x) = x^{\mu-1}(1+x)^{\nu-\mu}. \tag{2-2}$$

As $x \rightarrow 0$, $f_{\mu,\nu} = O(x^{\mu-1})$ and as $x \rightarrow \infty$, $f_{\mu,\nu} = O(x^{\nu-1})$. So the family may independently explore extreme behavior at both ends of the domain. Note that the existence of $\|f_{\mu,\nu}\|_p$ requires $p(\mu-1) > -1$ and $p(\nu-1) < -1$, equivalent to $\nu < 1/p' < \mu$.

We write

$$F_{\mu,\nu}(w) = \int_0^\infty \frac{f_{\mu,\nu}(x)}{(u+w)^\lambda} du \tag{2-3}$$

for the transform of $f_{\mu,\nu}$. However, the Lebesgue norm of this transform is analytically intractable, so at this stage we resort to numerical integration. A program has been constructed that computes the ratio

$$R = \frac{\|F_{\mu,\nu}\|_r}{\|f_{\mu,\nu}\|_p}. \tag{2-4}$$

It is used to numerically optimize this ratio as f ranges over the space of test functions, giving an estimate of

$$S = \sup_{\nu < 1/p' < \mu} \frac{\|F_{\mu,\nu}\|_r}{\|f_{\mu,\nu}\|_p}. \tag{2-5}$$

The optimization is repeated for a variety of values of p and λ and the result compared with Levin's constant $K(p, q)$, given in (1-3).

3. COMPUTATION OF THE FUNCTIONALS

For the test function (2-2) we have

$$\int_0^\infty f_{\mu,\nu}^p(u) du = \int_0^\infty u^{p(\mu-1)}(1+u)^{p(\nu-\mu)} du,$$

which is convergent at 0 and at ∞ if $\mu > 1/p'$ and $\nu < 1/p'$ respectively. The substitution $u = (1/s)-1$ converts this to a beta integral, showing that

$$\|f_{\mu,\nu}\|_p = B^{1/p}(p-p\nu-1, p\mu-p+1). \tag{3-1}$$

For the transform

$$F_{\mu,\nu}(w) = \int_0^\infty \frac{u^{\mu-1}(1+u)^{\nu-\mu}}{(u+w)^\lambda} du, \tag{3-2}$$

convergence requires $\mu > 0$ and $\nu < \lambda$. These will be satisfied if we again take $\mu > 1/p'$ and $\nu < 1/p'$, because

under the assumptions of Theorem 1.3, $0 < 1/p' < \lambda$. This integral may be obtained by conversion to the Euler integral

$$\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a} dt = B(b, c-b)G\left(\begin{matrix} a, b \\ c \end{matrix}; z\right),$$

where G is Gauss's hypergeometric function. (We use G rather than the conventional F or ${}_2F_1$ to avoid confusion with the left side of (1-6).) For $0 < w < 1$ the substitution required is $u = (1/t) - 1$, which gives

$$B(\lambda - \nu, \mu) G\left(\begin{matrix} \lambda - \nu, \lambda \\ \lambda + \mu + \nu \end{matrix}; 1 - w\right). \tag{3-3}$$

For $w > 1$ the substitution $u = w((1/t) - 1)$ yields

$$B(\lambda - \nu, \mu)w^{\nu-\lambda}G\left(\begin{matrix} \lambda - \nu, \mu - \nu \\ \lambda + \mu - \nu \end{matrix}; 1 - \frac{1}{w}\right). \tag{3-4}$$

These may be used to compute the norm of the transform,

$$\begin{aligned} \|F\|_r^r &= \int_0^\infty F^r(w) dw \\ &= \int_0^1 F^r(w) dw + \int_1^\infty F^r(w) dw = I_1 + I_2, \end{aligned} \tag{3-5}$$

where

$$I_1 = B^r(\lambda - \nu, \mu) \int_0^1 G^r\left(\begin{matrix} \lambda - \nu, \lambda \\ \lambda + \mu - \nu \end{matrix}; 1 - w\right) dw \tag{3-6}$$

and

$$\begin{aligned} I_2 &= B^r(\lambda - \nu, \mu) \\ &\times \int_1^\infty w^{r(\nu-\lambda)} G^r\left(\begin{matrix} \lambda - \nu, \mu - \nu \\ \lambda + \mu - \nu \end{matrix}; 1 - \frac{1}{w}\right) dw. \end{aligned} \tag{3-7}$$

It is these two integrals that are analytically intractable. We next discuss numerical approximations of them.

4. NUMERICAL EVALUATION

Consider first the improper integral I_1 above. From a programming perspective it is simpler to convert this to an infinite integral by inverting the variable of integration,

$$I_1 = B^r(\lambda - \nu, \mu) \int_1^\infty w^{-2} G^r\left(\begin{matrix} \lambda - \nu, \lambda \\ \lambda + \mu - \nu \end{matrix}; 1 - \frac{1}{w}\right) dw. \tag{4-1}$$

This is estimated by applying Simpson's rule to a finite approximation, \int_1^X for large X . The algorithm implemented accumulates terms of the form

$$\frac{h}{3}\{F(x) + 4F(x+h) + F(x+2h)\}.$$

We write H for the initial value of h . This is taken as 0.1, but since typically $X \approx 10^{50}$ is required to obtain reasonable accuracy, we need to increase h as the summation proceeds. This is achieved by doubling h when the sequence $\{F(x), F(x+h), F(x+2h)\}$ is close to linear, since then the intrinsic error in the Simpson method is small. More precisely, if

$$\left| \frac{F(x) + F(x+2h)}{2} - F(x+h) \right| < M,$$

then h is doubled at the next iteration. We found $M \approx 10^{-9}$ to be suitable. This reduces the number of iterations required to the order of 10^6 . (Selection of values for X and M is discussed further in Section 5.)

Note that each evaluation of the integrand in (4-1) requires a call to a function that evaluates the hypergeometric function G . This was coded by summing the series definition

$$G \left(\begin{matrix} a, b \\ c \end{matrix} ; x \right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{x^n}{n!}. \quad (4-2)$$

For w large in the integrand of (4-1), $1 - \frac{1}{w} \approx 1$, so the hypergeometric series (4-2) is very slowly convergent, making the computation impractical. However, if we use the analytic continuation of the hypergeometric function, [Erdélyi et al. 53],

$$\begin{aligned} G \left(\begin{matrix} a, b \\ c \end{matrix} ; x \right) & \quad (4-3) \\ &= \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} G \left(\begin{matrix} a, b \\ a+b-c+1 \end{matrix} ; 1-x \right) \\ &+ \frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)} (1-x)^{c-a-b} \\ &\times G \left(\begin{matrix} c-a, c-b \\ c-a-b+1 \end{matrix} ; 1-x \right), \end{aligned}$$

the resulting hypergeometric functions have argument close to 0 and so their computation is rapid. This was used as an alternative to direct summation of the series in computing the integrand of (4-1) when $1 - \frac{1}{w} > 0.6$.

Computation of I_2 is similar. Again this is speeded by applying the analytic continuation formula when $1 - 1/w > 0.6$.

A program `hlp_ratio` that computes these integrals and evaluates R for given parameters p , λ , μ , and ν was coded in C. The execution time of this program was about 15 seconds on a high-end Linux box when the integrals were rapidly convergent. But for the (often critical) cases in which the integrals were almost divergent, execution could take several minutes.

5. TESTING AND ERROR HANDLING

The approach of the previous section is subject to various sources of numerical error. These are considered in this section.

5.1 Rounding Error

All-floating point numbers are stored as 64-bit (17-digit) doubles. However, when one adds a large number of terms to estimate both the infinite integrals and the hypergeometric function, there is the potential for these errors to grow and swamp the results. As is well known, this is especially a problem if the terms are added in order of size, from the biggest to the smallest. To reduce the accumulated error we used a variant of Wolfe's method [Wolfe 64], with terms added in bins according to their order of magnitude.

To monitor the rounding errors a probabilistic analysis was performed, tracking the expected variance of the accumulated rounding error through the computation. It was found that the use of Wolfe's method reduced the expected standard deviation σ in R by several orders of magnitude. For a variety of cases investigated this standard deviation was less than R by a factor of 10^8 . For later experiments a check was inserted to abort the computation if $3\sigma/R > 0.01$.

5.2 Simpson's Rule Error

This is the error caused by using Simpson's rule to integrate a function that is not a cubic. This will be sensitive to the step size. Tests performed varying the initial step size H and the value of M that controls the growth of the step size showed that these affect only the eleventh significant figure in the result.

5.3 Truncation Error and Optimization Error

When the hypergeometric series (4-2) is summed, x never exceeds 0.6, so the series converges geometrically and the truncation error is negligible. For the infinite integrals, however, convergence may be very slow, especially when μ and ν are near $1/p'$, the limit of their ranges. Then, even with the greatest practicable value for X , 10^{150} , there is considerable error in the improper integrals (3-7) and (4-1).

For example, consider the case $p = 2$, $\lambda = 1$, in which the best constant is known to be π . With ν fixed and $\mu \rightarrow \frac{1}{2}$, then $R = \|F_{\mu,\nu}\|_r / \|f_{\mu,\nu}\|_p$ should approach π . Below we show values of R against μ with $\nu = -0.5$.

| | | | | |
|-------|--------|--------|--------|--------|
| μ | 0.525 | 0.520 | 0.515 | 0.510 |
| R | 3.0640 | 3.0703 | 3.0574 | 2.9756 |

There is a trade-off here between optimal μ and the truncation error. As $\mu \rightarrow \frac{1}{2}$, the convergence slows and the truncation errors increase. The closest result has an error exceeding 1%.

This problem is alleviated by using an estimate for the truncation errors. The tail of the integral for I_2 is

$$T_2 = B^r(\lambda - \nu, \mu) \times \int_X^\infty w^{r(\nu-\lambda)} G^r \left(\begin{matrix} \lambda - \nu, \mu - \nu \\ \lambda + \mu - \nu \end{matrix} ; 1 - \frac{1}{w} \right) dw. \quad (5-1)$$

Applying (4-3) converts the integral to

$$\int_X^\infty w^{r(\nu-\lambda)} \left[\frac{\Gamma(\lambda + \mu - \nu)\Gamma(\nu)}{\Gamma(\mu)\Gamma(\lambda)} G \left(\begin{matrix} \lambda - \nu, \mu - \nu \\ \nu + 1 \end{matrix} ; \frac{1}{w} \right) + \frac{\Gamma(\lambda + \mu - \nu)\Gamma(-\nu)}{\Gamma(\mu - \nu)\Gamma(\lambda - \nu)} w^{-\nu} G \left(\begin{matrix} \lambda, \mu \\ \nu + 1 \end{matrix} ; \frac{1}{w} \right) \right]^r dw. \quad (5-2)$$

Since the arguments $1/w$ are very small we may approximate hypergeometric functions by 1; selecting and integrating the larger of the two terms gives

$$T_2 \approx \begin{cases} \left[\frac{\Gamma(\mu)\Gamma(-\nu)}{\Gamma(\mu - \nu)} \right]^r \frac{X^{1-r\lambda}}{r\lambda - 1} & \text{if } \nu < 0, \\ \left[\frac{\Gamma(\nu)\Gamma(\lambda - \nu)}{\Gamma(\lambda)} \right]^r \frac{X^{1+r\nu-r\lambda}}{r\lambda - r\nu - 1} & \text{if } \nu > 0. \end{cases} \quad (5-3)$$

Similarly, the tail of I_1 is

$$T_1 \approx \begin{cases} \left[\frac{\Gamma(\mu)\Gamma(\lambda - \mu)}{\Gamma(\lambda)} \right]^r \frac{X^{r\lambda-r\mu-1}}{1 - r\lambda + r\mu} & \text{if } \lambda > \mu, \\ \left[\frac{\Gamma(\lambda - \nu)\Gamma(\mu - \lambda)}{\Gamma(\mu - \nu)} \right]^r \frac{1}{X} & \text{if } \lambda < \mu. \end{cases} \quad (5-4)$$

Including these estimates in the computation of the integrals (4-1) and (3-7) dramatically reduced truncation error in cases where it is known. For the case $p = 2$, $\lambda = 1$, $\nu = -0.5$ discussed above, optimal R now occurs at $\mu = 0.5000002$, and the optimum is 3.1415920, within 0.00002% of the exact optimum π . These results and others described in Section 8 give us confidence that our methodology yields meaningful results.

Occasionally, however, there was a problem with the tail estimate T_2 . When ν is small $\Gamma(\nu)$ is large and the two terms in (5-2) represent the difference of large and approximately equal numbers. The resulting numerical instability may give meaningless results. Consequently, to estimate the errors in (5-3) the alternative case is used. That is, for $\nu < 0$ we use the second case and for $\nu > 0$ the first. When the result indicates a possible error of more

than 1%, the computation is aborted. A similar check is applied to the use of (5-4) to catch problems when $\lambda \approx \mu$, although this did not occur with our experiments.

6. THE OPTIMIZATIONS

For a given $p > 1$, $\lambda \in (1/p', 1]$ we estimate (2-5) by performing a numerical optimization over $\mu > 1/p'$, $\nu < 1/p'$. We expect the optimum to occur near the boundaries of this region. So in practice we optimized over

$$1/p' + \epsilon \leq \mu \leq 1/p' + \epsilon + 2, \quad 1/p' + \epsilon - 2 \leq \nu \leq 1/p' + \epsilon,$$

where ϵ is a small positive number. We found $\epsilon = 0.001$ to be appropriate. The simplex algorithm [Nelder and Mead 65] was used with a starting point at the center of the search space.

Figure 1 shows the path of the optimal vertex for the case $p = 2$, $\lambda = 1$. The algorithm achieved a maximum at $\mu = 0.501$, $\nu = 0.499$, a corner of the space. The maximum attained, 3.14158, is close to the best constant π .

In order to find the dependence of the best constant on p and λ , optimizations like this must be performed for a range of values of p and λ . We used $\sqrt{\log_2 p} = \{0.1, 0.2, \dots, 2.0\}$ and $\lambda = \{0.05, 0.10, \dots, 1.0\} \cap (1/p', 1]$. This yields 209 optimizations, each of which requires typically about 50 evaluations, a considerable computational task. The following section addresses the method used to expedite this process.

7. THE NIMROD TOOLS

Nimrod/G [Abramson et al. 00] is a tool that expedites parametric computing. Given a computational model that requires input parameters, Nimrod/G allows the user to specify allowed values for each parameter, generates a job for each combination of these values, and executes these jobs on a cluster, or via the ‘‘computational grid,’’ on clusters and supercomputers around the world. The jobs will run in parallel with concurrency limited by the number of processors available. Normally a computational model needs no modification to run under Nimrod/G. The user prepares a text ‘‘plan’’ file that specifies the parameters and parameter values, and lists the tasks required for each combination.

To perform optimization on some aspect of the output of a computational model, Nimrod/O [Abramson et al. 01] is an appropriate tool. It offers a range of optimization algorithms. The implementation of these optimization procedures uses concurrent evaluations where

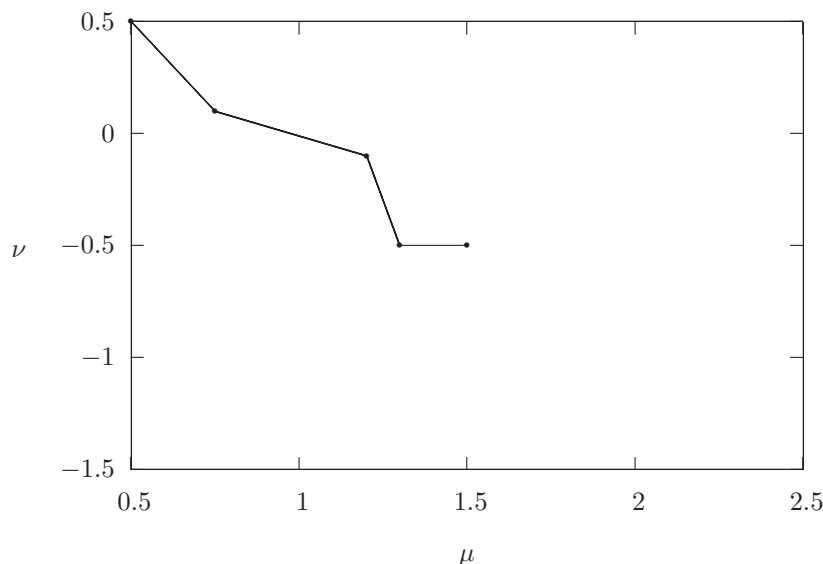


FIGURE 1. Track of (μ, ν) for optimization.

possible. In addition, it allows separate optimizations to be run concurrently.

A Nimrod/O optimization uses a “schedule” file. This is similar to a plan file with parameter specification and execution instructions. It also includes a section specifying the optimization method to use, the number of optimizations to run, and other settings relevant to the optimization algorithm.

We used Nimrod/O, with its built-in simplex algorithm, to perform each of the optimizations. Further, Nimrod/G was used to generate these optimizations, organizing the values of p and λ . This is a novel use of the tools, using one tool to launch multiple instances of the other, so we show below how this is done.

Figure 2 shows the plan file used to explore (2–5) for values of $\sqrt{\log_2 p}$ and λ . For each combination the Perl script `subst.pl` is executed. This computes the values of

```
parameter root_log_p float range from 0.1 to 2 step 0.1;
parameter lambda float range from 0.025 to 1 step 0.025;
task nodestart
  copy subst.pl node:.
  copy skeleton.shd node:.
  copy hlp_ratio node:.
  copy enfuzion.nodes node:.
endtask
task main
  node:execute ./subst.pl $root_log_p $lambda
  node:execute nimrodo -f hlp.shd -d s
  copy node:nimrodo-results.log
  results_$root_log_p.$lambda
endtask
```

FIGURE 2. Plan file for parameter sweep of values of p , λ .

p and p' , aborts if $\lambda \leq 1/p'$, and otherwise takes the file `skeleton.shd` and converts it to `hlp.shd`, a Nimrod/O schedule file tailored for this combination of p and λ . Nimrod/O is then executed and the results copied back to the root node. Figure 3 shows the file `hlp.shd` generated for the case $p = 2$, $\lambda = 1$.

```
parameter mu float range from 0.501 to 2.501
parameter nu float range from -1.501 to 0.499
task main
  copy hlp_ratio node:.
  node:execute ./hlp_ratio 2.0 1.0 $mu $nu
  copy node:output output.$jobname
endtask
method simplex
  starts 1
  tolerance 0.0001
  starting points specified
  (1.5 -0.5)
  endstarts
endmethod
```

FIGURE 3. Example schedule file for optimization over μ , ν .

8. RESULTS AND DISCUSSION

The optimizations were performed on a cluster of fifty 3.0-GHz processors running Linux. This required 10,602 runs of the program `hlp_ratio` with mean wall-clock execution time of 51 seconds. The mean time for an optimization was 24 minutes. The entire experiment completed in 9 hours and 12 minutes. This implies an average concurrency of 16.3.

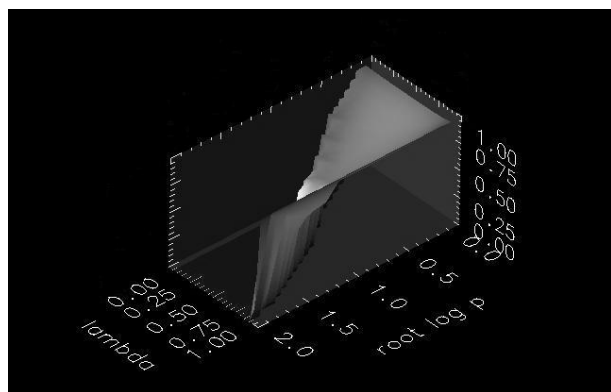


FIGURE 4. Surface showing S/K against $\sqrt{\log_2 p}$ and λ .

During some optimizations the execution of `hlp_ratio` failed due to the error estimate of T_2 as described above. In such cases Nimrod/O assigns the value -10^{100} and proceeds with the simplex optimization. In all cases a local maximum was found as indicated by a variation of $< 0.01\%$ over the vertices of the simplex.

Figure 4 shows a rubber-sheet diagram for S/K , the computed maximum transformation norm as a proportion of Levin’s constant, against $\sqrt{\log_2 p}$ and λ .¹ This was prepared using the graphics package OpenDX, which requires all values in the rectangular grid to be included. So for the forbidden region $\lambda \leq 1/p'$ we have assigned a value 0 to S/K .

The diagram shows that the results agree with known theory. For the (conjugate) case $\lambda = 1$ (the front right edge of the sheet) the ratio S/K is close to 1. Figures 5 and 6 show $\mu - 1/p'$ and $1/p' - \nu$ respectively versus $\sqrt{\log_2 p}$ and λ . With a few exceptions the optima for $\lambda = 1$ were obtained for μ and ν at the extremes of their allowed domains, namely $\frac{1}{p'} + 0.001$ and $\frac{1}{p'} - 0.001$.

Further, for $\sqrt{\log_2 p} = 0.1$, ($p = 2^{0.01} \approx 1$ at the back of the figure), S/K is again close to 1, agreeing with Walker’s result for $p \rightarrow 1$. Here there seems no obvious pattern in the μ and ν that attained the optima. Significantly, the optima did not occur at extreme values.

Elsewhere S/K is less than 1. Either the family of test functions (2-2) does not encompass the supremum (2-5) or the hypothesis that Levin’s constant is the best possible is wrong. Note that this family supplies only extreme behavior at 0 and ∞ . The only other family that we suggest may provide suitably extreme behavior

¹Thanks are due to Mr. Donny Kurniawan, who prepared the rubber-sheet diagrams.

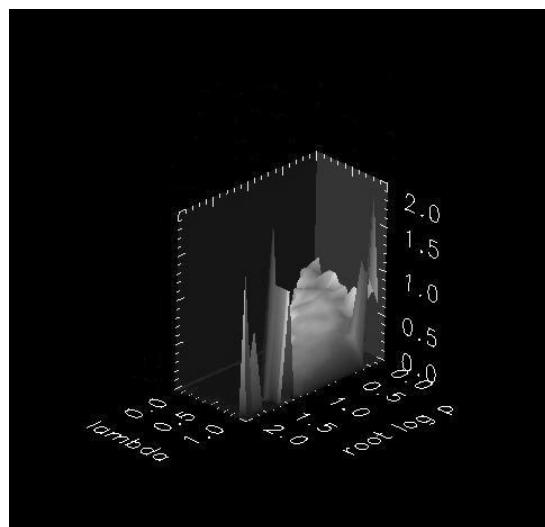


FIGURE 5. Surface showing optimal $\mu - 1/p'$ against $\sqrt{\log_2 p}$ and λ .

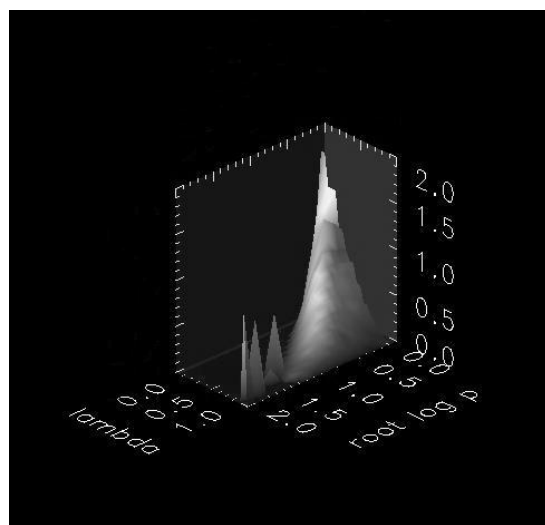


FIGURE 6. Surface showing optimal $1/p' - \nu$ against $\sqrt{\log_2 p}$ and λ .

is

$$f_{a,b}(x) = \begin{cases} b^{-1/p} & \text{for } a - b/2 \leq x \leq a + b/2, \\ 0 & \text{otherwise.} \end{cases}$$

This has $\|f_{a,b}\| = 1$ and “approaches” a Dirac delta at $x = a$ as $b \rightarrow 0$.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the Victorian Partnership for Advanced Computing.

REFERENCES

- [Abramson et al. 00] D. Abramson, J. Giddy, and L. Kotler. “High Performance Parametric Modeling with Nimrod/G: Killer Application for the Global Grid?” In *International Parallel and Distributed Processing Symposium (IPDPS)*, May 2000.
- [Abramson et al. 01] D. Abramson, A. Lewis, and T. C. Peachey. “An Automatic Design Optimization Tool and Its Application to Computational Fluid Dynamics.” In *Supercomputing 2001*, July 2001.
- [Bonsall 51] F. F. Bonsall. “Inequalities with Non-conjugate Parameters.” *Quart. J. Math.* 2 (1951), 135–150.
- [Erdélyi et al. 53] A. Erdélyi et al. *Higher Transcendental Functions*. Vol. 1. New York: McGraw Hill, 1953.
- [Finch 03] S. R. Finch. *Mathematical Constants*. Cambridge: Cambridge University Press, 2003.
- [Hardy et al. 52] G. H. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities*. Second edition. Cambridge: Cambridge University Press, 1952.
- [Levin 36] V. Levin. “On the Two Parameter Extension and Analogue of Hilbert’s Inequality.” *J. of London Math. Soc.* 11 (1936), 119–124.
- [Nelder and Mead 65] J. A. Nelder and R. Mead. “A Simplex Method for Function Minimization.” *Computer Journal* 7 (1965), 308–313.
- [Peachey 03] T. C. Peachey. “Some Integral Inequalities Related to Hilbert’s.” *Journal of Inequalities in Pure and Applied Mathematics* 4:1 (2003).
- [Peachey et al. 99] T. C. Peachey, A. McAndrew, and S. S. Dragomir. “The Best Constant in an Inequality of Ostrowski Type.” *Tamkang Journal of Mathematics* 30:3 (1999), 219–222.
- [Walker 73] P. L. Walker. “A Note on an Inequality with Non-conjugate Parameters.” *Proceedings of the Edinburgh Mathematical Society* 18 (1973), 293–294.
- [Wolfe 64] J. M. Wolfe. “Reducing Truncation Errors by Programming.” *Comm. of the ACM* 7 (1964), 355–356.

T. C. Peachey, Caulfield School of Information Technology, Monash University, Caulfield East, Australia, 3145
(tcp@csse.monash.edu.au)

C. M. Enticott, CRC for Enterprise Distributed Systems Technology, c/o Caulfield School of Information Technology,
Monash University, Caulfield East, Australia, 3145 (cme@csse.monash.edu.au)

Received November 11, 2004; accepted July 22, 2005.