In *Information Processing Letters* **45** (1993) 41–50.

A Systematic Analysis of Splaying

Berry Schoenmakers
CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands
berry@cwi.nl

Abstract

In this paper we perform an amortized analysis of a functional program for splaying. We construct a potential function that yields the same bound for the amortized cost of splaying as given by D.D. Sleator and R.E. Tarjan—the inventors of splay trees. In addition, we show that this bound is minimal for the class of "sum of logs" potentials. Our approach also applies to the analysis of path reversal and pairing heaps.

Keywords Analysis of algorithms, data structures, amortized complexity, potential function, functional programming, splay trees.

1 Splaying

Splaying is the central operation in a particular implementation of *dictionaries*. A dictionary is an abstract data type involving operations on subsets of an infinite, linearly ordered set, such as the integers. In [5] Sleator and Tarjan developed an efficient implementation of dictionaries, called *splay trees*. Splay trees are *binary search trees*, i.e. binary trees of integers whose inorder traversal is *strictly increasing*. There is no balance condition imposed on these trees whatsoever, but it is the particular way splaying is defined that makes the data structure efficient.

Operation $a\angle$ ("splaying at a") is performed by "rotating" a to the root of a binary search tree while keeping the inorder traversal intact (see Figure 1). In cases (iii) and (iv), triangle y stands for the upper part of the tree that remains unchanged. The symmetrical counterparts of (ii)–(iv), which are obtained by interchanging the role of left and right subtrees, are omitted. Transformations (iii) and (iv) are repeatedly applied until (i) or (ii) applies. It is assumed that a occurs in the tree.

We translate this pictorial description of splaying into the following functional program, in which $\langle t, a, u\rangle$ denotes a nonempty binary tree with left subtree t, root a, and right subtree u (again, the symmetrical cases are omitted):

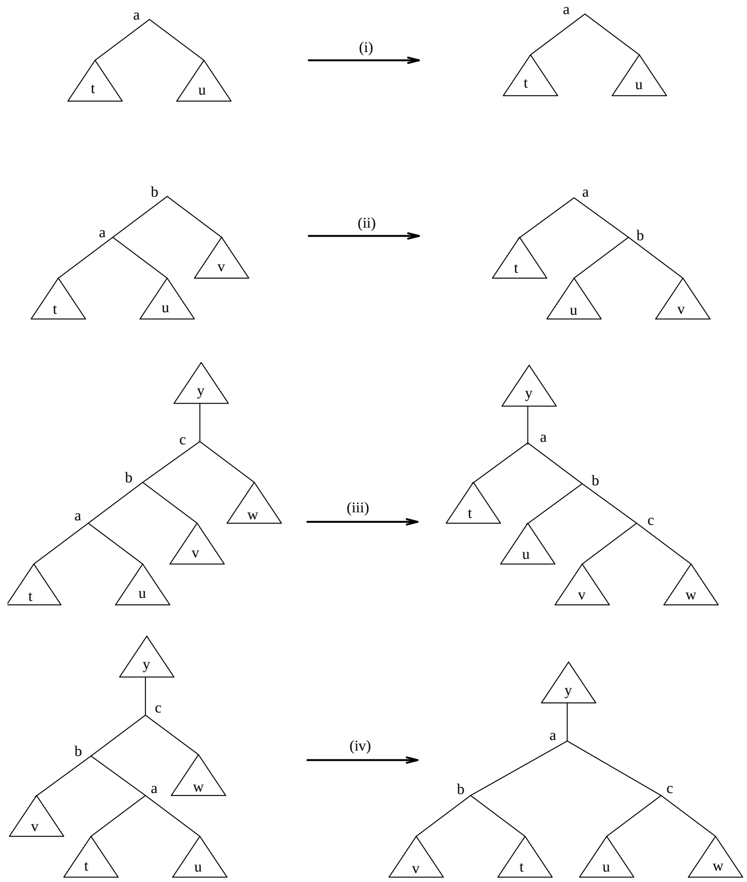

Figure 1: Splaying at a (cf. [5, Figure 3]).

$$
\begin{array}{lllll}
\text{(i)} & a\angle\,\langle t,a,u\rangle & = & \langle t,a,u\rangle & \\
\text{(ii)} & a\angle\,\langle\langle t,a,u\rangle,b,v\rangle & = & \langle t,a,\langle u,b,v\rangle\rangle & ,a<b \\
\text{(iii)} & a\angle\,\langle\langle x,b,v\rangle,c,w\rangle & = & \langle t,a,\langle u,b,\langle v,c,w\rangle\rangle\rangle \ \text{ where } \langle t,a,u\rangle = a\angle\,x & ,a<b<c \\
\text{(iv)} & a\angle\,\langle\langle v,b,x\rangle,c,w\rangle & = & \langle\langle v,b,t\rangle,a,\langle u,c,w\rangle\rangle \ \text{ where } \langle t,a,u\rangle = a\angle\,x & ,b<a<c.
\end{array}
$$

Clearly, the root of $a\angle\,x$ equals a, and the inorder traversals of x and $a\angle\,x$ are equal. If the depth of a is even, the program returns the same tree as the operation described by Figure 1. In case the depth of a is odd, the resulting trees may be quite different. However, as far as efficiency is concerned, the difference is not essential, since both versions of splaying can be analyzed in the same way.

2 Analysis

Evaluation of $a\angle$ amounts to repeatedly unfolding either (iii) or (iv), followed by a single unfolding of either (i) or (ii). A useful cost measure is therefore given by $T.x =$ "number of unfoldings of (iii) and (iv) required for the evaluation of $a\angle\,x$." Given cost measure T, we want to derive a logarithmic bound for the amortized cost A of $a\angle$, given by

$$A.x = T.x + \Phi.(a\angle\,x) - \Phi.x,$$

where Φ is a potential function [6]. For the sake of brevity it it left implicit that $T.x$ and $A.x$ depend on a.

Setting out for an inductive derivation, we first calculate a recurrence relation for A. To this end, we take Φ of the form:

$$\begin{array}{lcl} \Phi.\langle\,\rangle & = & 0 \\ \Phi.\langle t,a,u\rangle & = & \Phi.t + \varphi.t.u + \Phi.u. \end{array}$$

Note that a does not occur in the right-hand side of the definition of $\Phi.\langle t,a,u\rangle$; this reflects our decision to let $\Phi.x$ depend on the structure of x only. Moreover, $\varphi.t.u$ will be defined to be symmetric in t and u, so that the symmetrical counterparts of (ii)–(iv) can be ignored in the sequel. Now, we calculate for case (iv):

$$\begin{array}{ll}
 & A.\langle\langle v,b,x\rangle,c,w\rangle \\
= & \{\text{ definition of } A \text{ }\} \\
 & T.\langle\langle v,b,x\rangle,c,w\rangle + \Phi.(a\angle\,\langle\langle v,b,x\rangle,c,w\rangle) - \Phi.\langle\langle v,b,x\rangle,c,w\rangle \\
= & \{\text{ definitions of } a\angle \text{ and } T \text{ }\} \\
 & 1 + T.x + \Phi.\langle\langle v,b,t\rangle,a,\langle u,c,w\rangle\rangle - \Phi.\langle\langle v,b,x\rangle,c,w\rangle \\
= & \{\text{ definition of } A;\ a\angle\,x = \langle t,a,u\rangle \text{ }\} \\
 & 1 + A.x + \Phi.x - \Phi.\langle t,a,u\rangle + \Phi.\langle\langle v,b,t\rangle,a,\langle u,c,w\rangle\rangle - \Phi.\langle\langle v,b,x\rangle,c,w\rangle \\
= & \{\text{ definition of } \Phi \text{ }\} \\
 & 1 + A.x + \Phi.x - \Phi.t - \varphi.t.u - \Phi.u \\
 & +\Phi.v + \varphi.v.t + \Phi.t + \varphi.\langle v,b,t\rangle.\langle u,c,w\rangle + \Phi.u + \varphi.u.w + \Phi.w \\
 & -\Phi.v - \varphi.v.x - \Phi.x - \varphi.\langle v,b,x\rangle.w - \Phi.w \\
= & \{\text{ simplifying }\} \\
 & A.x + 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u \\
 & +\varphi.\langle v,b,t\rangle.\langle u,c,w\rangle - \varphi.\langle v,b,x\rangle.w \\
= & \{\text{ see below }\} \\
 & A.x + 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u.
\end{array}$$

Let $|x|$ denote one plus the size of x: $|\langle\,\rangle| = 1$ and $|\langle t,a,u\rangle| = |t| + |u|$. Then, for the last step, note that $|t| + |u| = |x|$ because $\langle t,a,u\rangle = a\angle\,x$ and $|a\angle\,x| = |x|$. Hence, $\varphi.\langle v,b,t\rangle.\langle u,c,w\rangle = \varphi.\langle v,b,x\rangle.w$ provided $\varphi.t.u$ depends on $|t| + |u|$ only. This proviso will be assumed in the sequel. Similar calculations for cases (ii) and (iii) then yield as recurrence relation for A:

$$\begin{array}{lcl}
A.\langle t,a,u\rangle & = & 0 \\
A.\langle\langle t,a,u\rangle,b,v\rangle & = & \varphi.u.v - \varphi.t.u \\
A.\langle\langle x,b,v\rangle,c,w\rangle & = & A.x + 1 + \varphi.u.\langle v,c,w\rangle + \varphi.v.w - \varphi.x.v - \varphi.t.u \\
A.\langle\langle v,b,x\rangle,c,w\rangle & = & A.x + 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u.
\end{array}$$

A logarithmic bound on A follows if we are able to define φ such that, for instance,

(1) $\quad A.x \le \log_\alpha |x|$

with $\alpha > 1$. Inspired by [5, Lemma 1], however, and in view of Theorem 1 (Section 3), we will prove the following stronger bound:

(2) $\quad A.x \le \log_\alpha \dfrac{|x|}{|x|_a}$.

Here $|x|_a$ denotes one plus the size of a's subtree in x.

In case (i), (2) evidently holds. In order that (2) follows by induction in the other cases, the following requirements are imposed on φ:

$$\begin{array}{lcl} \varphi.u.v - \varphi.t.u & \le & \log_\alpha \dfrac{|t|+|u|+|v|}{|t|+|u|} \\ 1 + \varphi.u.\langle v,c,w\rangle + \varphi.v.w - \varphi.x.v - \varphi.t.u & \le & \log_\alpha \dfrac{|t|+|u|+|v|+|w|}{|t|+|u|} \\ 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u & \le & \log_\alpha \dfrac{|t|+|u|+|v|+|w|}{|t|+|u|}. \end{array}$$

The last requirement corresponds to case (iv), and results from the following calculation:

$$\begin{array}{ll} & A.\langle\langle v,b,x\rangle,c,w\rangle \\ = & \{ \text{ above recurrence relation } \} \\ & A.x + 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u \\ \le & \{ \text{ induction hypothesis (2) } \} \\ & \log_\alpha \dfrac{|x|}{|x|_a} + 1 + \varphi.v.t + \varphi.u.w - \varphi.v.x - \varphi.t.u \\ \le & \{ \text{ last requirement on } \varphi;\ |x| = |t| + |u| \text{ and } |x|_a = |\langle\langle v,b,x\rangle,c,w\rangle|_a \ \} \\ & \log_\alpha \dfrac{|\langle\langle v,b,x\rangle,c,w\rangle|}{|\langle\langle v,b,x\rangle,c,w\rangle|_a}. \end{array}$$

The important observation is now that these requirements are not only sufficient but also necessary for (2) to hold. This is clear in case (ii), since in this case the requirement on φ is just a reformulation of (2). To see this for the other two requirements, we reason as follows. Consider case (iv) and take $x = \langle t,a,u\rangle$. Then $A.x = 0$ and $|x| = |x|_a = |\langle\langle v,b,x\rangle,c,w\rangle|_a = |t| + |u|$, and—as may be gathered from the above calculation—the last requirement on φ is then *equivalent* to (2) in this case. The same reasoning applies to case (iii).

Next, to remove $\log_\alpha$ from the above requirements, we define $\varphi.t.u$ as the following function of $|t| + |u|$:

$$\varphi.t.u = \beta \log_\alpha(|t| + |u|),$$

with $\beta \neq 0$. On account of the monotonicity of $\log_\alpha$ for $\alpha > 1$, the requirements on φ then reduce to the following requirements on α and β:

$$\begin{array}{rcl} \left(\dfrac{|u|+|v|}{|t|+|u|}\right)^\beta & \le & 1 + \dfrac{|v|}{|t|+|u|} \\ \alpha\left(\dfrac{|u|+|v|+|w|}{|t|+|u|+|v|}\,\dfrac{|v|+|w|}{|t|+|u|}\right)^\beta & \le & 1 + \dfrac{|v|+|w|}{|t|+|u|} \\ \alpha\left(\dfrac{|v|+|t|}{|v|+|t|+|u|}\,\dfrac{|u|+|w|}{|t|+|u|}\right)^\beta & \le & 1 + \dfrac{|v|+|w|}{|t|+|u|}. \end{array}$$

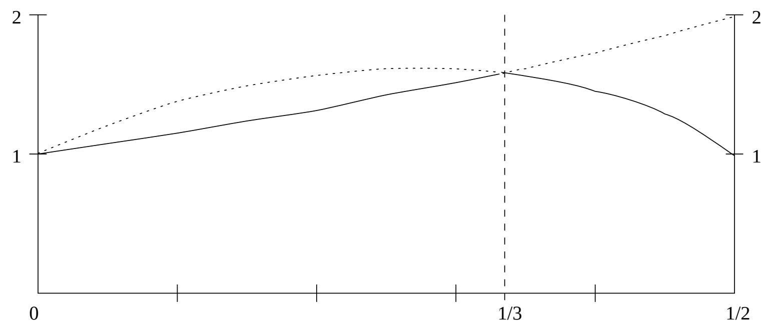

Figure 2: Maximum at $\beta = \frac{1}{3}$.

Summarizing, we have as sufficient and necessary constraints on α and β to guarantee (2):

(3) $\quad (\forall\, k, l, m :: (\dfrac{l+m}{k+l})^{\beta} \leq 1 + \dfrac{m}{k+l})$

(4) $\quad (\forall\, k, l, m, n :: \alpha(\dfrac{l+m+n}{k+l+m}\ \dfrac{m+n}{k+l})^{\beta} \leq 1 + \dfrac{m+n}{k+l})$

(5) $\quad (\forall\, k, l, m, n :: \alpha(\dfrac{k+m}{k+l+m}\ \dfrac{l+n}{k+l})^{\beta} \leq 1 + \dfrac{m+n}{k+l})$.

Under these constraints, we now maximize α so as to minimize bound (2). We distinguish three cases, using that $\alpha > 1$ and $\beta \neq 0$.

Case $\beta < 0$. Instantiation of (3) with $l, m := 1, 1$ yields that $(\frac{2}{k+1})^{\beta} \leq 1 + \frac{1}{k+1}$ for all k. But this is false if $\beta < 0$.

Case $0 < \beta < \frac{1}{2}$. In the appendix it is proved that in this case:

$$\begin{array}{lcll} (3) & \equiv & \text{true} & \text{(Lemma 1)} \\ (4) & \equiv & \alpha \leq \frac{(1-\beta)^{1-\beta}}{\beta^{\beta}(1-2\beta)^{1-2\beta}} & \text{(Lemma 2)} \\ (5) & \equiv & \alpha \leq 4^{\beta} & \text{(Lemma 3)}\ , \end{array}$$

so, taking the conjunction of the requirements on α and β, we obtain:

$$1 < \alpha \leq \frac{(1-\beta)^{1-\beta}}{\beta^{\beta}(1-2\beta)^{1-2\beta}} \min 4^{\beta}.$$

To maximize α, we determine the maximum of its upper bound over all β satisfying $0 < \beta < \frac{1}{2}$. This yields $\alpha = \sqrt[3]{4}$ (≈ 1.59) as maximal value at $\beta = \frac{1}{3}$ (see also Figure 2).

Case $\beta \geq \frac{1}{2}$. Instantiation of (4) with $k, l, m := 1, 1, 1$ yields that $\alpha(\frac{n+2}{3}\ \frac{n+1}{2})^{\beta} \leq \frac{n+3}{2}$, for all n, which is equivalent to $\alpha \leq \frac{6^{\beta}}{2} \frac{n+3}{(n^2+3n+2)^{\beta}}$. This upper bound is decreasing in β, so to minimize it we take $\beta = \frac{1}{2}$. Taking $n \to \infty$, we then get $\alpha \leq \sqrt{\frac{3}{2}}$.

Since $\sqrt{\frac{3}{2}} < \sqrt[3]{4}$, we conclude from this case analysis that the maximal value for α is given by $\alpha = \sqrt[3]{4}$ (for $\beta = \frac{1}{3}$).

3 Result

The analysis in the previous section constitutes a proof of the following theorem.

Theorem 1 Let potential Φ be of the form

$$\begin{array}{lcl} \Phi.\langle\,\rangle & = & 0 \\ \Phi.\langle t,a,u\rangle & = & \Phi.t + \beta \log_\alpha |\langle t,a,u\rangle| + \Phi.u \end{array}$$

with $\alpha > 1$ and $\beta \neq 0$. Then A satisfies

$$(\forall\, a,x :: A.x \leq \log_\alpha \frac{|x|}{|x|_a}) \;\equiv\; (3) \;\wedge\; (4) \;\wedge\; (5),$$

and this bound on A is minimal for $\alpha = \sqrt[3]{4}$ (and $\beta = \frac{1}{3}$). □

In other words, $\frac{3}{2}\log_2(|x|/|x|_a)$ is the minimal bound for $A.x$ for the class of "sum of logs" potentials. A better bound for $A.x$ of this form can only be obtained by using a different type of potential function.

To obtain the bound of Sleator and Tarjan [5, Lemma 1], we define $T'.x = 2 + 2\,T.x$. Then $T'.x$ corresponds to the cost measure in [5], which counts the number of integer comparisons required for the evaluation of $a\angle\, x$. Furthermore, we define $A'.x = T'.x + \Phi'.(a\angle\, x) - \Phi'.x$ where $\Phi'.x = 2\,\Phi.x$. Since $A'.x = 2 + 2\,A.x$, Theorem 1 now yields

$$A'.x \leq 2 + 3\log_2 \frac{|x|}{|x|_a} \leq 2 + 3\log_2 |x|$$

as bound on the amortized number of comparisons used by splaying. The corresponding potential is given by

$$\begin{array}{lcl} \Phi'.\langle\,\rangle & = & 0 \\ \Phi'.\langle t,a,u\rangle & = & \Phi'.t + \log_2 |\langle t,a,u\rangle| + \Phi'.u. \end{array}$$

4 Concluding remarks

In a systematic way, we have analyzed a top-down version of splaying. We have shown that using (2) (or (1)) as induction hypothesis leads relatively straightforward to a "sum of logs" potential. The corresponding bound on the amortized costs matches the bound of Sleator and Tarjan. By carefully deriving requirements on Φ that are sufficient and necessary for (2) to hold, we showed in addition that this bound is minimal for the class of "sum of logs" potentials.

Along the same lines, path reversal [2] and the two-pass variant of pairing [1] can be analyzed (see [4]). In all these analyses a "sum of logs" potential arises. As the authors of [2] note, this is easy to explain for splaying and pairing, since there is a clear connection between these operations (see [1, p.121]). But they are at a loss to explain that such a potential can also be used to amortize the cost of path reversal, because they cannot discover a connection between splaying (and pairing) on the one hand, and path reversal on the other. Our analyses, however, show that "sum of logs" potentials can be derived systematically as one sets out to prove logarithmic bounds for the amortized costs of these operations. Another example of such a derivation is the analysis of top-down skew heaps in [3], in which an asymmetric variant of a "sum of logs" potential arises.

Acknowledgements The referees are gratefully acknowledged for their useful comments which improved the presentation.

References

1. Fredman M.L., Sedgewick R., Sleator D.D., Tarjan R.E. *The pairing heap: a new form of self-adjusting heap.* Algorithmica 1 (1986) 111–129.
2. Ginat D., Sleator D.D., Tarjan R.E. *A tight amortized bound for path reversal.* Information Processing Letters 31 (1989) 3–5.
3. Kaldewaij A., Schoenmakers B. *The derivation of a tighter bound for top-down skew heaps.* Information Processing Letters 37 (1991) 265–271.
4. Schoenmakers B. *Data Structures and Amortized Complexity in a Functional Setting.* Ph.D. thesis, Eindhoven University of Technology, Eindhoven, Netherlands (1992).
5. Sleator D.D., Tarjan R.E. *Self-adjusting binary search trees.* Journal of the ACM 32 (1985) 652–686.
6. Tarjan R.E. *Amortized computational complexity.* SIAM Journal on Algebraic and Discrete Methods 6 (1985) 306–318.

Appendix

Lemma 1 For $\beta > 0$,

$$(\forall\, k, l, m :: \left(\frac{l{+}m}{k{+}l}\right)^{\beta} \leq 1 + \frac{m}{k{+}l}) \;\equiv\; \beta \leq 1.$$

Proof By mutual implication:

$$\begin{array}{ll} & (\forall\, k, l, m :: (\tfrac{l+m}{k+l})^{\beta} \leq 1 + \tfrac{m}{k+l}) \\ \Leftarrow & \{\ \beta > 0\ \} \\ & (\forall\, k, l, m :: (\tfrac{k+l+m}{k+l})^{\beta} \leq \tfrac{k+l+m}{k+l}) \\ \equiv & \{\ \text{algebra}\ \} \\ & \beta \leq 1; \end{array}$$

$(\forall k,l,m :: (\frac{l+m}{k+l})^\beta \le 1 + \frac{m}{k+l})$

$\Rightarrow$ { take $k = 1$ and $l = 1$ }

$(\forall m :: (\frac{m+1}{2})^\beta \le \frac{m+2}{2})$

$\Rightarrow$ { take $m \to \infty$ }

$\beta \le 1$.

□

Lemma 2 For $0 < \beta < \frac{1}{2}$,

$$(\forall k,l,m,n :: \alpha \left(\frac{l+m+n}{k+l+m}\ \frac{m+n}{k+l}\right)^\beta \le 1 + \frac{m+n}{k+l}) \ \equiv\ \alpha \le \frac{(1-\beta)^{1-\beta}}{\beta^\beta(1-2\beta)^{1-2\beta}}.$$

Proof By mutual implication:

$(\forall k,l,m,n :: \alpha(\frac{l+m+n}{k+l+m}\ \frac{m+n}{k+l})^\beta \le 1 + \frac{m+n}{k+l})$

$\Leftarrow$ { $\beta > 0$ }

$(\forall k,l,m,n :: \alpha(\frac{k+l+m+n}{k+l}\ \frac{m+n}{k+l})^\beta \le 1 + \frac{m+n}{k+l})$

$\equiv$ { $p = \frac{m+n}{k+l}$ }

$(\forall p : p > 0 : \alpha((1+p)p)^\beta \le 1 + p)$

$\equiv$ { simplifying }

$(\forall p : p > 0 : \alpha \le \frac{(1+p)^{1-\beta}}{p^\beta})$

$\equiv$ { minimize $\frac{(1+x)^{1-\beta}}{x^\beta}$, using $0 < \beta < \frac{1}{2}$ (see below) }

$\alpha \le \frac{(1-\beta)^{1-\beta}}{\beta^\beta(1-2\beta)^{1-2\beta}}$;

$(\forall k,l,m,n :: \alpha(\frac{l+m+n}{k+l+m}\ \frac{m+n}{k+l})^\beta \le 1 + \frac{m+n}{k+l})$

$\Rightarrow$ { take $k = m = 1$ }

$(\forall l,n :: \alpha(\frac{l+n+1}{l+2}\ \frac{n+1}{l+1})^\beta \le 1 + \frac{n+1}{l+1})$

$\equiv$ { algebra }

$(\forall l,n :: \alpha \le (1 + \frac{n+1}{l+1})(\frac{l+2}{l+n+1}\ \frac{l+1}{n+1})^\beta)$

$\Rightarrow$ { take $\frac{n}{l} \to \frac{\beta}{1-2\beta}$ and $l \to \infty$ ($\frac{\beta}{1-2\beta} > 0$, since $0 < \beta < \frac{1}{2}$) }

$\alpha \le (1 + \frac{\beta}{1-2\beta})(\frac{1}{1+\frac{\beta}{1-2\beta}}\ \frac{1-2\beta}{\beta})^\beta$

$\equiv$ { $1 + \frac{\beta}{1-2\beta} = \frac{1-\beta}{1-2\beta}$ }

$\alpha \le \frac{1-\beta}{1-2\beta}(\frac{1-2\beta}{1-\beta}\ \frac{1-2\beta}{\beta})^\beta$

$\equiv$ { simplifying }

$\alpha \le \frac{(1-\beta)^{1-\beta}}{\beta^\beta(1-2\beta)^{1-2\beta}}$.

To complete the proof we minimize $f(x) = \frac{(1+x)^{1-\beta}}{x^\beta}$ over all $x > 0$, for $0 < \beta < \frac{1}{2}$. Then $f'(x) = 0$ is equivalent to $(1-\beta)x = (1+x)\beta$, and f turns out to be minimal at $\frac{\beta}{1-2\beta}$, which is positive because $0 < \beta < \frac{1}{2}$. So the minimum of f equals $\frac{(1+\frac{\beta}{1-2\beta})^{1-\beta}}{(\frac{\beta}{1-2\beta})^\beta}$, which in turn equals $\frac{(1-\beta)^{1-\beta}}{\beta^\beta(1-2\beta)^{1-2\beta}}$, as $1 + \frac{\beta}{1-2\beta} = \frac{1-\beta}{1-2\beta}$. □

Lemma 3 For $0 < \beta < \frac{1}{2}$,

$$(\forall\, k,l,m,n :: \alpha \left(\frac{k+m}{k+l+m}\ \frac{l+n}{k+l}\right)^\beta \leq 1 + \frac{m+n}{k+l}) \ \equiv\ \alpha \leq 4^\beta.$$

Proof By mutual implication:

$(\forall\, k,l,m,n :: \alpha(\frac{k+m}{k+l+m}\ \frac{l+n}{k+l})^\beta \leq 1 + \frac{m+n}{k+l})$

$\Leftarrow$ { $\beta > 0$ }

$(\forall\, k,l,m,n :: \alpha(\frac{k+m}{k+l}\ \frac{l+n}{k+l})^\beta \leq 1 + \frac{m+n}{k+l})$

$\equiv$ { $p = \frac{k+m}{k+l}$ and $q = \frac{l+n}{k+l}$ }

$(\forall\, p,q : p > 0 \ \wedge\ q > 0 \ \wedge\ p+q \geq 1 : \alpha(pq)^\beta \leq p+q)$

$\equiv$ { algebra }

$(\forall\, p,q : p > 0 \ \wedge\ q > 0 \ \wedge\ p+q \geq 1 : \alpha \leq \frac{p+q}{(pq)^\beta})$

$\equiv$ { minimize $\frac{x+y}{(xy)^\beta}$, using $0 < \beta < \frac{1}{2}$ (see below) }

$\alpha \leq 4^\beta$;

$(\forall\, k,l,m,n :: \alpha(\frac{k+m}{k+l+m}\ \frac{l+n}{k+l})^\beta \leq 1 + \frac{m+n}{k+l})$

$\Rightarrow$ { take $l = k$ and $m = n = 1$ }

$(\forall\, k :: \alpha(\frac{k+1}{2k+1}\ \frac{k+1}{2k})^\beta \leq 1 + \frac{2}{2k})$

$\equiv$ { simplifying }

$(\forall\, k :: \alpha \leq (1+\frac{1}{k})(\frac{2k+1}{k+1}\ \frac{2k}{k+1})^\beta)$

$\Rightarrow$ { take $k \to \infty$ }

$\alpha \leq 4^\beta$.

We determine the minimum of $\frac{x+y}{(xy)^\beta}$ over all positive x and y satisfying $x+y \geq 1$, when $0 < \beta < \frac{1}{2}$. To this end, we first observe that this function takes on its minimal value only if $x = y$ because $(xy)^\beta$ is maximized by taking x equal to y ($\beta > 0$), and this can be achieved without changing the value of $x+y$. We thus minimize $2x^{1-2\beta}$ over all $x \geq \frac{1}{2}$. Since this function is increasing in x ($\beta < \frac{1}{2}$), it attains its minimal value of 4^β at $x = \frac{1}{2}$. □