We first define the function h and put the number of digits high, to have better computer precision

$$> \quad h(x) := -\frac{(x\cdot\log(x) + (1-x)\cdot\log(1-x))}{\log(2)};$$

$$h(0) := 0;$$

$$h(1) := 0;$$

$$h := x \mapsto -\frac{x\cdot\log(x) + (1-x)\cdot\log(1-x)}{\log(2)}$$

$$h(0) := 0$$

$$h(1) := 0 \qquad \textbf{(1)}$$

The number of digits can be adjusted. (it has been done with larger accuracy as well, but some outcomes become unconveniently lengthy)

$$> \quad Digits := 20;$$

$$Digits := 20 \qquad \textbf{(2)}$$

We compute the values for b,a and c which are related with the atomic distribution which shows that the improvement is bounded (by c).

$$> \quad with(RootFinding):$$

$$B := NextZero\left(x \mapsto h(x) * (2-h(x)) - h\left(2*x-x^2\right), 0.14\right);$$

$$B := 0.32945473850303697239 \qquad \textbf{(3)}$$

$$> \quad evalf[10](B);$$

$$0.3294547385 \qquad \textbf{(4)}$$

$$> \quad A := \frac{(1-h(B))}{(2-h(B))};$$

$$A := \frac{1-\dfrac{0.63379181713619122525}{\ln(2)}}{2-\dfrac{0.63379181713619122525}{\ln(2)}} \qquad \textbf{(5)}$$

$$> \quad evalf[10](A);$$

$$0.07887729268 \qquad \textbf{(6)}$$

$$> \quad C := A + (1-A)\cdot B;$$

$$C := \frac{0.67054526149696302761\left(1-\dfrac{0.63379181713619122525}{\ln(2)}\right)}{2-\dfrac{0.63379181713619122525}{\ln(2)}} + 0.32945473850303697239 \qquad \textbf{(7)}$$

$$> \quad evalf[10](C);$$

$$0.3823455334 \qquad \textbf{(8)}$$

Next, we consider the ratio's of the derivatives to compute the optimal choice of alpha.

$$> \quad g1(x) := diff\left(((1-(C-x)/(1-x))\text{^}2*h\left(2*x-x^2\right)-(1-(C-x)/(1-x))*h(x), x\right);$$

$$g1 := x \rightarrow \frac{\partial}{\partial x}\left(\left(1\,\mathrm{K}\ \frac{C\,\mathrm{K}\ x}{1\,\mathrm{K}\ x}\right)^2 h\left(2\,x\,\mathrm{K}\ x^2\right)\mathrm{K}\ \left(1\,\mathrm{K}\ \frac{C\,\mathrm{K}\ x}{1\,\mathrm{K}\ x}\right) h(x)\right) \quad \textbf{(9)}$$

> $g2(x) := diff\left(\left(1\mathrm{K}\ 2*(C\mathrm{K}\ x)/(1\mathrm{K}\ x)\right)\mathrm{K}\ \left(1\mathrm{K}\ (C\mathrm{K}\ x)/(1\mathrm{K}\ x)\right)*h(x), x\right);$

$$g2 := x \rightarrow \frac{\partial}{\partial x}\left(1\,\mathrm{K}\ \frac{2\,C\,\mathrm{K}\ 2\,x}{1\,\mathrm{K}\ x}\,\mathrm{K}\ \left(1\,\mathrm{K}\ \frac{C\,\mathrm{K}\ x}{1\,\mathrm{K}\ x}\right) h(x)\right) \quad \textbf{(10)}$$

> $\text{alpha} := \mathrm{K}\ \dfrac{eval(g1(x), x=B)}{(eval(g2(x), x=B)\mathrm{K}\ eval(g1(x), x=B))};$

$$\alpha := \mathrm{K}\ \Bigg(\frac{1}{\ln(2)}\Bigg(1.3761289611414332582\Bigg(1 \quad \textbf{(11)}$$

$$\mathrm{K}\ \frac{1.0000000000000000000\left(1\,\mathrm{K}\ \dfrac{0.63379181713619122525}{\ln(2)}\right)}{2\,\mathrm{K}\ \dfrac{0.63379181713619122525}{\ln(2)}}\Bigg)$$

$$\Bigg(1.4913236397607875851$$

$$\mathrm{K}\ \frac{1.4913236397607875851\left(1\,\mathrm{K}\ \dfrac{0.63379181713619122525}{\ln(2)}\right)}{2\,\mathrm{K}\ \dfrac{0.63379181713619122525}{\ln(2)}}\Bigg)\Bigg)$$

$$\mathrm{K}\ \frac{1}{\ln(2)}\Bigg(0.27111744475595941547\Bigg(1$$

$$\left. K \frac{1.0000000000000000000 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}} \right)^{2} \right)$$

$$K \frac{1}{\ln(2)} \left(0.63379181713619122525 \left(1.4913236397607875851 \right.\right.$$

$$\left.\left. K \frac{1.4913236397607875851 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}} \right) \right)$$

$$K \frac{1}{\ln(2)} \left(0.71065222421202958606 \left(1 \right.\right.$$

$$\left.\left.\left. K \frac{1.0000000000000000000 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}} \right) \right) \right) \Bigg/$$

$$\left(2.9826472795215751702\,K \frac{2.9826472795215751702 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}} \right.$$

$$K \frac{1}{\ln(2)} \left(1.3761289611414332582 \left(1 \right.\right.$$

$$K \frac{1.0000000000000000000 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}}\Bigg)$$

$$\Bigg(1.491236397607875851$$

$$K \frac{1.491236397607875851 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}}\Bigg)\Bigg)$$

$$+ \frac{1}{\ln(2)} \Bigg(0.27111744475595941547 \Bigg(1$$

$$K \frac{1.0000000000000000000 \left(1\,K \frac{0.63379181713619122525}{\ln(2)}\right)}{2\,K \frac{0.63379181713619122525}{\ln(2)}}\Bigg)^2\Bigg)\Bigg)$$

> $evalf[10](\mathrm{alpha})$

$$0.03560698066 \qquad \textbf{(12)}$$

We verify the statement for probability distributions for which the support contains 3 elements; a_1,a_2 and 1, which have probabilities respectively equal to p_1, p_2 and 1-p_1-p_2. As such we can compute E [H(p)], E[H(p+q-pq)] and E'[H(min(2*p,1/2)]

>

> $H(p\_1, a\_1, p\_2, a\_2) := p\_1 \cdot h(a\_1) + p\_2 \cdot h(a\_2);$

$$H := (p\_1, a\_1, p\_2, a\_2) \mapsto p\_1 \cdot h(a\_1) + p\_2 \cdot h(a\_2) \qquad \textbf{(13)}$$

> $Hpq(p\_1, a\_1, p\_2, a\_2) := p\_1^2 \cdot h\left((1\,K\ a\_1)^2\right) + p\_2^2 \cdot h\left((1\,K\ a\_2)^2\right) + 2 \cdot p\_1 \cdot p\_2 \cdot h((1\ K\ a\_1) \cdot (1\,K\ a\_2));$

$$Hpq := (p\_1, a\_1, p\_2, a\_2) \mapsto p\_1^2 \cdot h\left((1\,K\ a\_1)^2\right) + p\_2^2 \cdot h\left((1\,K\ a\_2)^2\right) + 2 \cdot p\_1 \cdot p\_2 \cdot h((1\,K\ a\_1) \cdot (1\,K\ a\_2)) \qquad \textbf{(14)}$$

> $Hpr(p\_1, a\_1, p\_2, a\_2) := (p\_1 + \min((p\_2\,K\ (1\,K\ p\_1\,K\ p\_2)), 0)) \cdot h\left(\min\left(2 \cdot a\_1, \frac{1}{2}\right)\right) + \max((p\_2\,K\ (1\,K\ p\_1\,K\ p\_2)), 0) \cdot h\left(\min\left(2 \cdot a\_2, \frac{1}{2}\right)\right);$

(15)

$$Hpr := (p\_1, a\_1, p\_2, a\_2) \mapsto (p\_1 + \min(2 \cdot p\_2 \,\mathrm{K}\, 1 + p\_1, 0)) \cdot h\left(\min\left(2 \cdot a\_1, \frac{1}{2}\right)\right) + \max(2 \cdot p\_2 \,\mathrm{K}\, 1 + p\_1, 0) \cdot h\left(\min\left(2 \cdot a\_2, \frac{1}{2}\right)\right) \quad \textbf{(15)}$$

We now consider the optimization problem for the two considered cases;
hereby we solve the minimization problem with two different substitutions (once with E[H(p)] <= C and once with E[H(p)] = C)
and also plot the 2D graph to see that the inequality is indeed true for the mentioned value C.

Case 1: (a_1,a_2,1) getting weights (1-2p_2,p_2.p_2), In this case the difference is zero and attained by the earlier mentioned atomic distribution since a_1=a_2)

> $with(Optimization):$
> $Minimize((1 \,\mathrm{K}\, \text{alpha}) \cdot Hpq(1 \mathrm{K}\, 2\, p\_2, a\_1, p\_2, a\_2) + \text{alpha} \cdot Hpr(1 \mathrm{K}\, 2\, p\_2, a\_1, p\_2, a\_2) \,\mathrm{K}\, H(1 \mathrm{K}\, 2\, p\_2, a\_1, p\_2, a\_2), \{p\_2 \le 1, a\_1 \le C, a\_2 \le 0.5, (1 \,\mathrm{K}\, 2\, p\_2) \cdot a\_1 + p\_2 \cdot a\_2 + p\_2 \le C, a\_1 \le a\_2\}, assume = nonnegative)$

$$[\mathrm{K}\, 4.801 \times 10^{\mathrm{K}\,17}, [a\_1 = 0.32945473850294784466, a\_2 = 0.32945473850294784466, p\_2 = 0.078877292706045659566]] \quad \textbf{(16)}$$

> $F(p\_1, a\_1, p\_2, a\_2) := (1 \,\mathrm{K}\, \text{alpha}) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \text{alpha} \cdot Hpr(p\_1, a\_1, p\_2, a\_2) \,\mathrm{K}\, H(p\_1, a\_1, p\_2, a\_2)$

$$F := (p\_1, a\_1, p\_2, a\_2) \mapsto (1 \,\mathrm{K}\, \alpha) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \alpha \cdot Hpr(p\_1, a\_1, p\_2, a\_2) \,\mathrm{K}\, H(p\_1, a\_1, p\_2, a\_2) \quad \textbf{(17)}$$

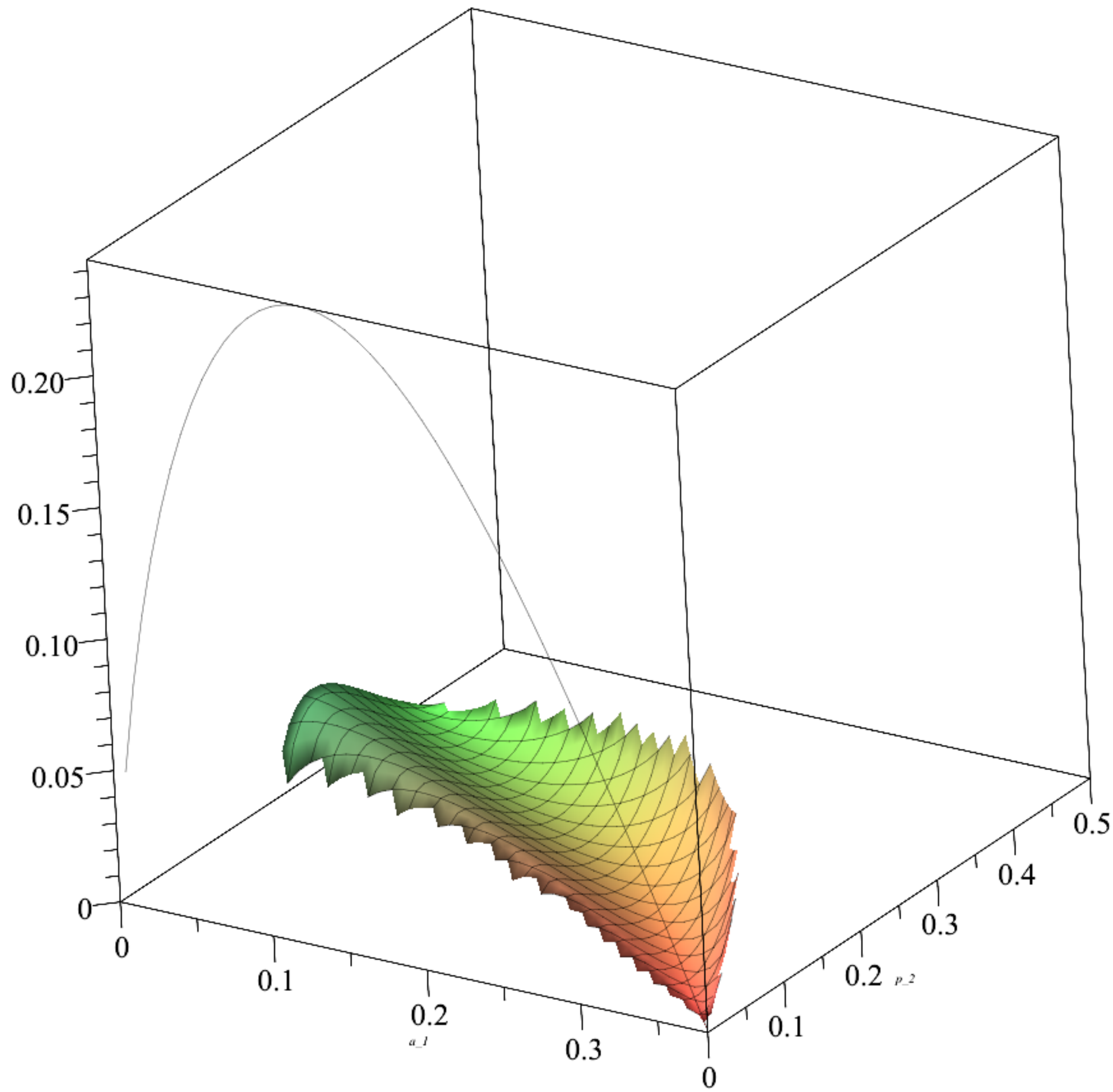
> $with(Optimization):$
> $Minimize\left(F\left(1 \,\mathrm{K}\, 2 \cdot \frac{(C \,\mathrm{K}\, a\_1)}{(1 + a\_2 \,\mathrm{K}\, 2 \cdot a\_1)}, a\_1, \frac{(C \,\mathrm{K}\, a\_1)}{(1 + a\_2 \,\mathrm{K}\, 2 \cdot a\_1)}, a\_2\right), \{0 \le a\_1, a\_1 \le C, a\_1 \le a\_2, a\_2 \le 0.5\}\right);$

$$[\mathrm{K}\, 1.8 \times 10^{\mathrm{K}\,20}, [a\_1 = 0.32945473850303617469, a\_2 = 0.32945473850303807804]] \quad \textbf{(18)}$$

>

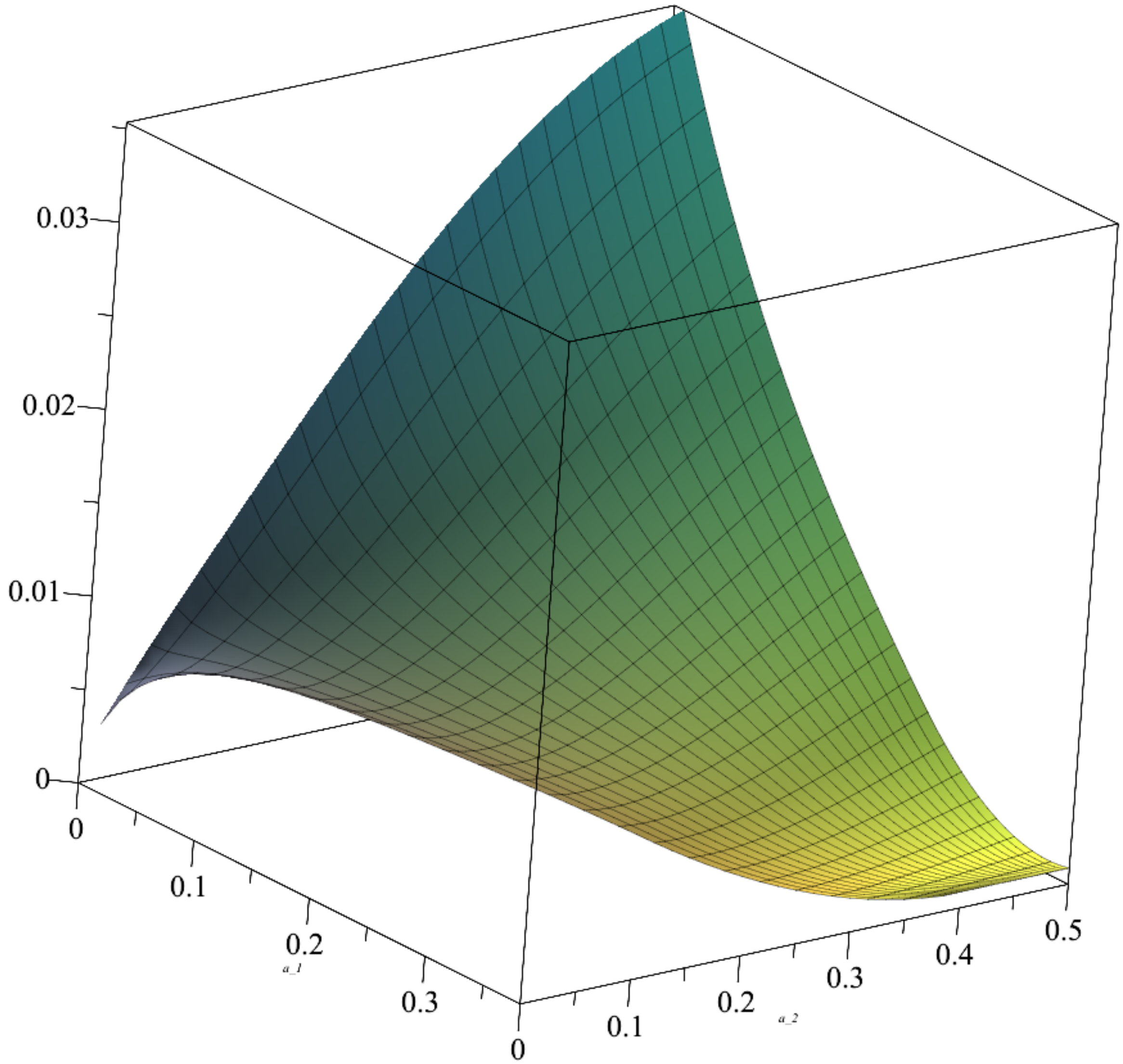
>

> $plot3d\left(F\left(1 \,\mathrm{K}\, 2\, p\_2, a\_1, p\_2, \frac{(C \,\mathrm{K}\, p\_2 \,\mathrm{K}\, (1 \,\mathrm{K}\, 2 \cdot p\_2) \cdot a\_1)}{p\_2}\right), a\_1 = 0 \,..\, C, p\_2 = 0 \,..\, 0.5\right);$

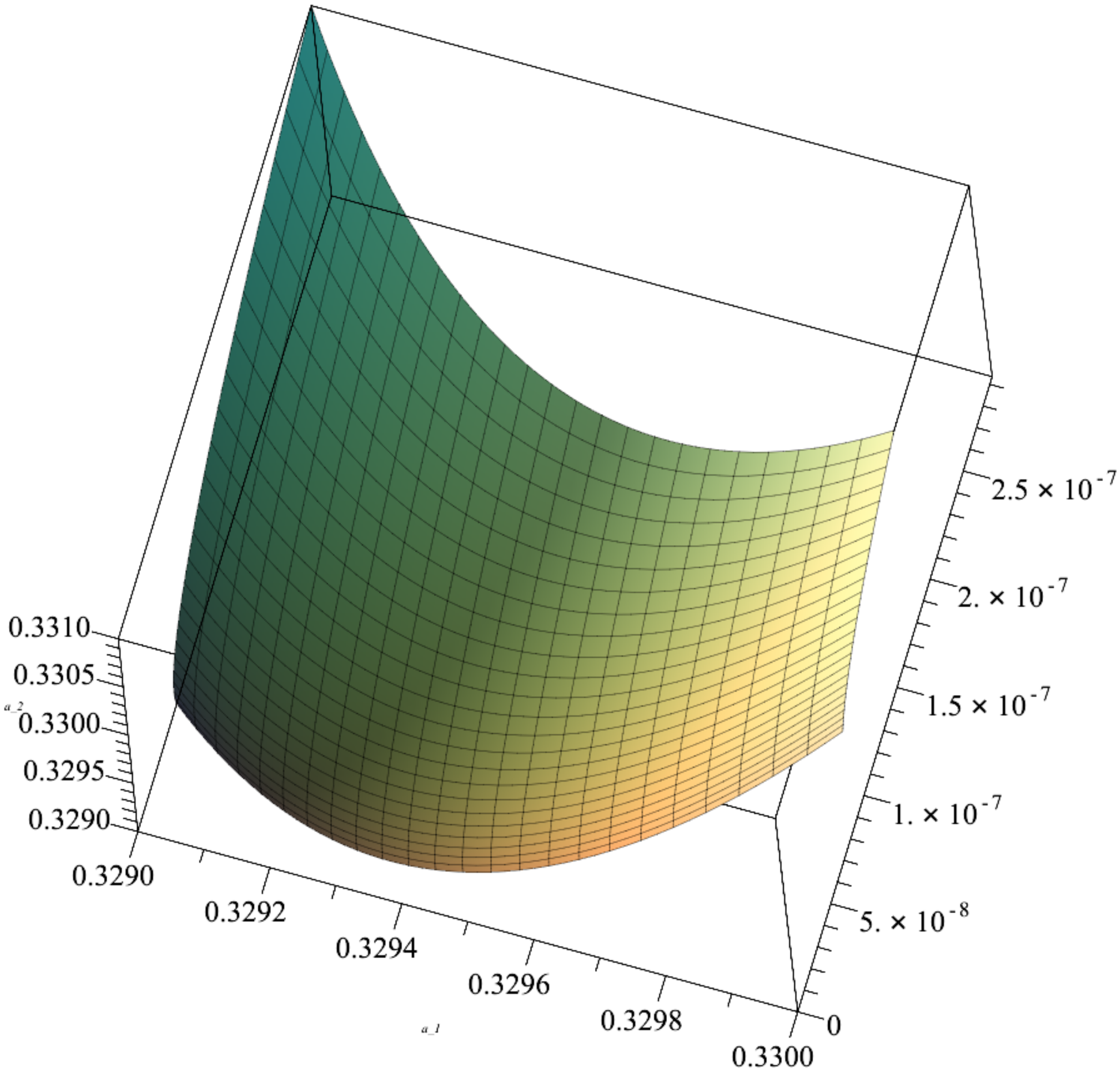


(19)

> $plot3d\left(F\left(1 \,\mathrm{K}\, 2\cdot\frac{(C\,\mathrm{K}\; a\_1)}{(1 + a\_2\,\mathrm{K}\; 2\cdot a\_1)}, a\_1, \frac{(C\,\mathrm{K}\; a\_1)}{(1 + a\_2\,\mathrm{K}\; 2\cdot a\_1)}, a\_2\right), a\_1 = 0\; ..C, a\_2 = a\_1\; ..\; 0.5\right);$



> $plot3d\left(F\left(1\,K\;2\cdot\frac{(C\,K\;a\_1)}{(1+a\_2\,K\;2\cdot a\_1)},a\_1,\frac{(C\,K\;a\_1)}{(1+a\_2\,K\;2\cdot a\_1)},a\_2\right),a\_1=0.329\,..0.33,\right.$
$\left.a\_2=a\_1\,..\,0.331\right);$



Case 2: (a_1,a_2) getting weights (1-p_2,p_2),In this case the difference is even strict. (the plot shows that the initial point is indeed a global maximum)

> $with(Optimization):$
$Minimize(\,(1\,\mathrm{K}\ \text{alpha})\cdot Hpq(1\,\mathrm{K}\ p\_2, a\_1, p\_2, a\_2) + \text{alpha}\cdot\ Hpr(1\,\mathrm{K}\ p\_2, a\_1, p\_2, a\_2)\ \mathrm{K}\ H(1\,\mathrm{K}\ p\_2, a\_1, p\_2, a\_2), \{\, p\_2 \leq 1, a\_1 \leq C, a\_2 \leq 0.5, (1\,\mathrm{K}\ p\_2)\cdot a\_1 + p\_2\cdot a\_2 \leq C, a\_1 \leq a\_2\}, assume = nonnegative)\ ;$

Warning, no iterations performed as initial point satisfies first-order conditions

$$[0.000867301752543841\,97, [a\_1 = 0.38234553336670272114, a\_2 = 0.38234553336670272114, p\_2 = 1.0]] \quad \textbf{(20)}$$

\>

\> $with(Optimization):$

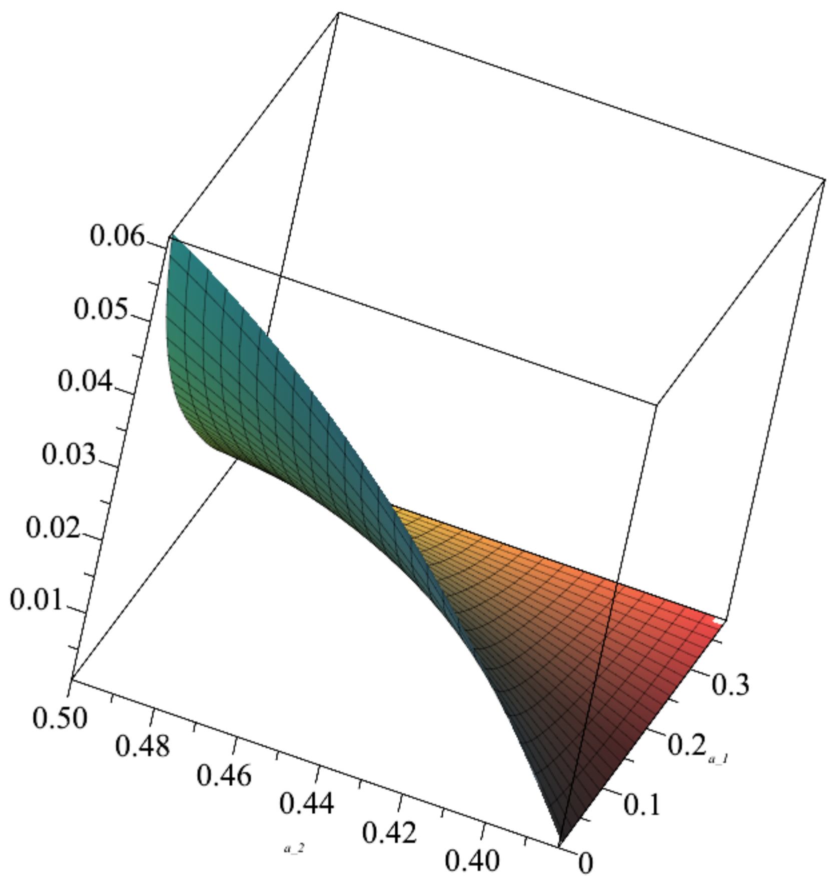
$Minimize\left(F\left(1 \mathrm{K}\ \frac{(C\mathrm{K}\ a\_1)}{(a\_2\mathrm{K}\ a\_1)}, a\_1, \frac{(C\mathrm{K}\ a\_1)}{(a\_2\mathrm{K}\ a\_1)}, a\_2\right), \{0 \le a\_1, a\_1 \le C, C \le a\_2, a\_2 \le 0.5\}\right);$

Warning, no iterations performed as initial point satisfies first-order conditions

$[0.000867301752543841\,93, [a\_1 = 0.38234553336670272114, a\_2 = 0.50000000000000000000]]$ **(21)**

\>

\> $plot3d\left(F\left(1 \mathrm{K}\ \frac{(C\mathrm{K}\ a\_1)}{(a\_2\mathrm{K}\ a\_1)}, a\_1, \frac{(C\mathrm{K}\ a\_1)}{(a\_2\mathrm{K}\ a\_1)}, a\_2\right), a\_1 = 0\,..C, a\_2 = C\,..\,0.5\right);$



\>

\>

\>