多倍長計算手法

平成25年度第2四半期

目次

１．はじめに

平成２５年第１四半期では
量子モンテカルロ法による物性スペクトル計算
で正しく計算できる範囲はDD形式では表現できる
数値範囲の制限により，β＝27,U=10までと
いう事が明白になった。
このため今四半期はieee754-2008にデータ
形式を使用して正しく計算出来るパラメータ範囲
を更に広げた。結果の検証は静的かつ規則的な
パスに対してG（β）＝－１でチェックしている。

またループ積分においてDE
（二重指数関数型積分法）の誤差の評価法が
固まったので今回記載しました。

2。 グリーン関数計算（続き）

前四半期は

$dt \times u < 1$　　$fac = \sqrt{dt \times u}$　　条件数 $= e^{fac \times \frac{l}{2}}$

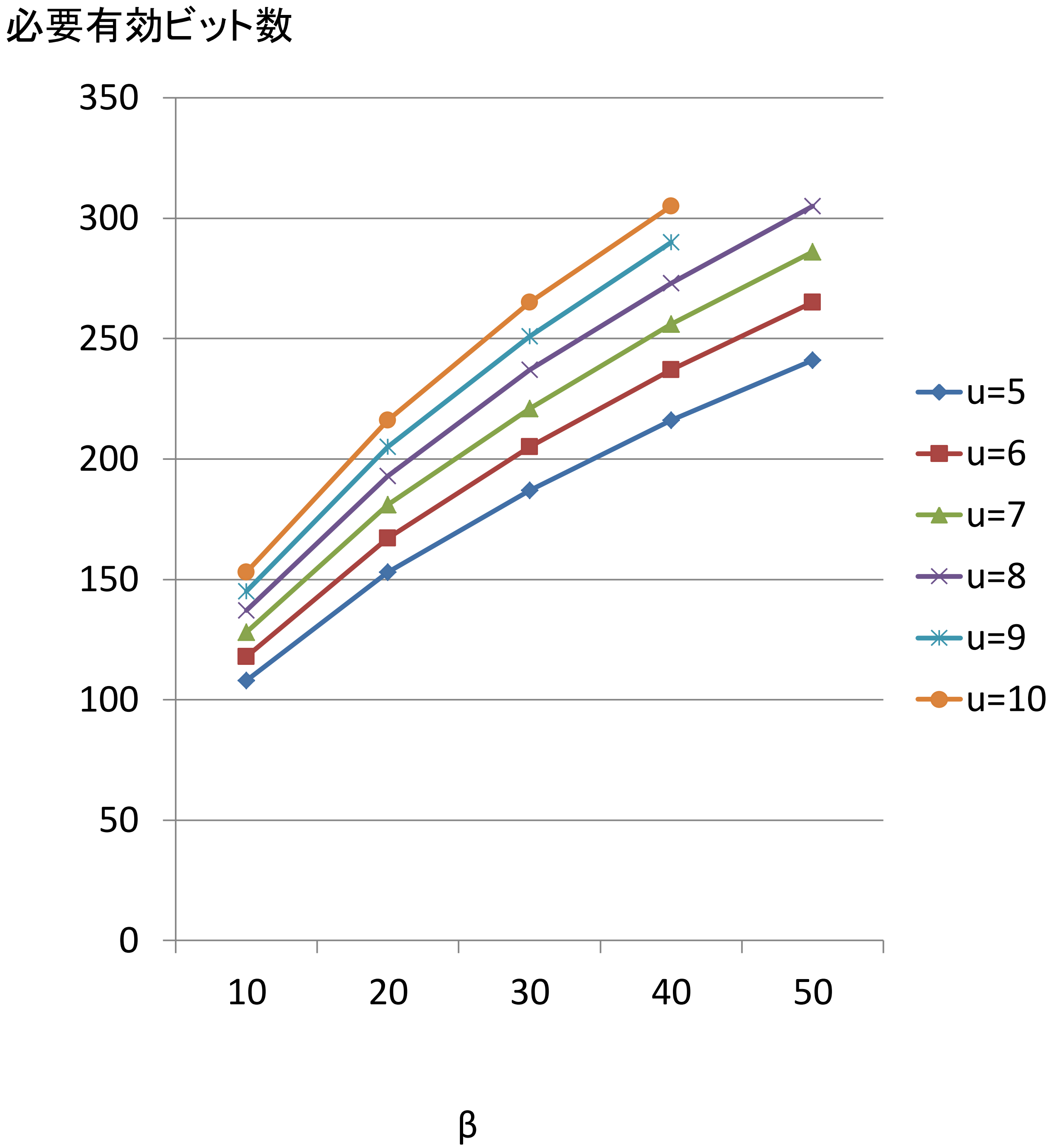
の $\beta = 30, u = 10$ までのケースを扱っていましたが今回は

$dt \times u > 1, dt \leq 1$　　条件数 $= e^{(dt \times u)^{dt} \times \frac{l}{2}}$

のケースや $dt \times u > 1, dt > 1$ での場合を記述しました。

$$dt = \frac{\beta}{l}, dt \times u < 1, dt < 1 \text{で} l = 448 \text{における必要有効ビット数}$$

dd:106　td:159 qd:212 fd:265
q:113　dq:226　tq:339 qq:452 fq:565



ieee754-2008 DQ形式結果一覧

l	β	u	精度	結果
448	20	9	DQ	OK
448	30	10	TQ	OK
448	40	10	TQ	OK
448	50	5	TQ	OK
448	50	6	TQ	OK
448	50	7	TQ	OK
448	50	8	TQ	OK
448	50	8	TQ	OK
448	50	9	TQ	4桁一致
448	50	10	16bai	OK
448	50	9	QQ	OK
448	50	10	QQ	OK
448	50	10	FQ	OK

32倍精度限界テスト

ieee形式限界テスト			
	演算精度	32倍精度	l=448,n=100
	beta	u	結果
	150	10	OK
	170	10	OK
	180	10	OK
	190	10	OK
	200	10	OK
	250	10	OK
	300	5	OK
	300	6	OK
	300	7	OK
	300	8	OK
	300	9	NG
	300	10	NG
	400	3	OK
	400	4	OK
	400	5	NG
	500	1	OK
	500	1.5	OK
	500	1.6	OK
	500	1.7	OK
	500	1.8	OK
	500	1.9	NG(8桁一致)
	500	2	NG
	500	3	NG

β＝500の場合はdt＞1と今までと少し様子が異なる.

β=500,u=1の場合

$dt>1$の場合,$x=dt$とすると,打ち切り誤算は$O(x^2)=e^x-x-1$となり,xよりちいさくないと正しい結果は得られない.

$e^x-x-1=x$となるxは$x=1.256431209$.

$limit=1.256431209$とおくと

$dt\times u>1, 1<dt<limit$ で

$$条件数=e^{(dt\times u)^{\sqrt{dt\times u}}\times\frac{\beta+l}{4}}$$となる.

$dt>limit$の場合正しい結果はだせない.

$l=448$ だと

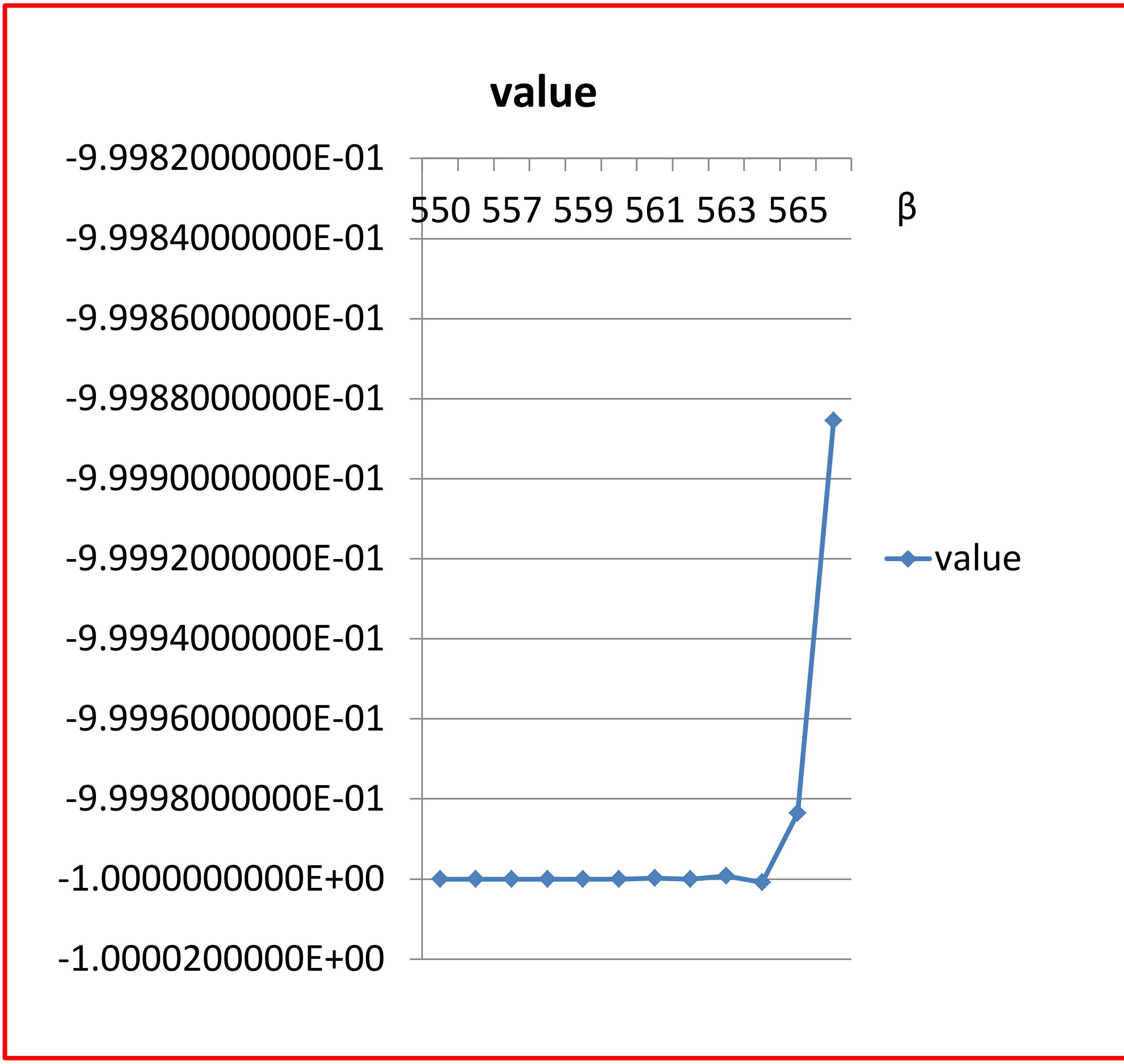
$\beta=limit\times l=562.88\fallingdotseq562$程度まで.

beta	u	value
520	1	−0.1000000000D+01
540	1	−0.1000000000D+01
550	1	−0.1000000000D+01
556	1	−0.1000000000D+01
557	1	−0.1000000000D+01
558	1	−0.1000000000D+01
559	1	−0.1000000000D+01
560	1	−0.9999999999D+00
561	1	−0.9999998502D+00
562	1	−0.1000000000D+01
563	1	−0.9999992758D+00
564	1	−0.1000000846D+01
565	1	−0.9999835911D+00
566	1	−0.9998855364D+00
568	1	−0.1448403101D+01
570	1	−0.2199689781D+01
576	1	0.2759895707D+09
578	1	−0.4289705228D+10
580	1	−0.1575586965D+09
590	1	−0.6396738371D+21
600	1	0.8569813910D+32

β=562.88まで正しく計算出来るとの見積もりとよく一致しています。

U=1　での実行結果一覧表

(L=448,32倍精度演算)



幾つかの特殊なケース

$$ieee754-2008 \quad 32倍精度演算$$
$$n=40, L=112$$

β	u	結果
50	200	OK
100	100	OK
200	50	NG
400	25	NG
1000	10	NG

lの値を小さくすると良くなるケース

演算精度	beta	u	l	結果
8倍精度	50	10	448	NG
8倍精度	50	10	400	NG
8倍精度	50	10	300	OK
16倍精度	100	10	300	OK
32倍精度	250	10	448	OK
32倍精度	300	10	300	OK

別解析と対称性を仮定した計算

$O(x^2)=e^x-x-1=x$となるxをx_0とする.

$x_0=1.256431208626192229633814710501 14$

$l=448$なら$l\times x_0=562.88$

$x=dt,u=1$とすると

$$e^{-\frac{x}{2}}\left(\frac{e^{\sqrt{x}}+e^{-\sqrt{x}}}{2}\right)\geq x^2x_0$$

$dt=\frac{\beta}{l}$とし,$l=896$とすると

$\beta\leq779$の場合正しく計算でき,$\beta\geq780$で正しく計算できない。

実測結果は以下のとうり。

$\beta=700$ OK

$\beta=720$ OK

$\beta=740$ OK

$\beta=760$ OK

$\beta=780$ NG

$\beta=780$ の場合$dt=776.5256125$まで正しい値となっている.

$\beta=800$ NG

$\beta=800$ の場合$dt=778.5714286$まで正しい値となっている.

$O(x^2)=e^x-x-1=x$となるxをx_0とする.

$x_0=1.25643120862619222963381471050114$

$x=dt=\dfrac{\beta}{l}$とすると$\beta\le lx_0$.また$e^l\le 2^m$ (m:有効ビット数)

これより, $\dfrac{\beta}{x_0}\le l\le m\times\ln(2)$

$m=1009$なら$\beta=800, l=694$はこれを満足する.$l=700$はぎりぎり.

k番目の固有値をa_kとすると$a_{n-k}=-a_k\ (k=1,2,....,\dfrac{n}{2})$より

$$S=\sum_{k=1}^{n}\frac{e^{-ta_k}}{1+e^{-\beta a_k}}=\sum_{k=1}^{\frac{n}{2}}\left(\frac{e^{-ta_k}}{1+e^{-\beta a_k}}+\frac{e^{ta_k}}{1+e^{\beta a_k}}\right)$$

$$=\sum_{k=1}^{\frac{n}{2}}\left(\frac{e^{-ta_k}+e^{(\beta-t)a_k}+e^{ta_k}+e^{(t-\beta)a_k}}{(1+e^{-\beta a_k})(1+e^{\beta a_k})}\right)$$

分母の値はβによらない。$t=\dfrac{\beta}{2}-m*dt$とする。$(0\le m\le\dfrac{l}{2})$

分子$=e^{(-\frac{\beta}{2}+m*dt)a_k}+e^{(\frac{\beta}{2}+m*dt)a_k}+e^{(\frac{\beta}{2}-m*dt)a_k}+e^{(-\frac{\beta}{2}-m*dt)a_k}$

となりmのところを$-m$にしてもSは不変なので, $\dfrac{\beta}{2}$に対して対称となる.

よって, Sは$m=\dfrac{l}{2}$で極値を取り, 最小値となる.

$O(x^2)=e^x-x-1=x$ となる x を x_0 とする.

$x_0=1.2564312086261922296338147105011 4$

$x=\sqrt{l\beta}$ 　　　$l<\beta\le x_0 l$ とする.

このときの平方根計算の引数 x には以下の制限が加わる。

m: 有効ビット数

$$2^{[m-\frac{1}{2}(\frac{x}{\ln 2}-m)]\sqrt{x_0}}\ge e^x$$

これから, $(\frac{3}{2}m-\frac{x}{2\ln 2})\sqrt{x_0}\ln 2\ge x$

$$x\le\frac{\frac{3}{2}m\ln 2\sqrt{x_0}}{1+\frac{\sqrt{x_0}}{2}}$$

$m=1009$ だと $x\le 753.5750183$

			理論	結果
$l=694$	$\beta=816$	$x=752.5317269$	OK	OK
$l=694$	$\beta=820$	$x=754.3739126$	NG	NG
$l=700$	$\beta=808$	$x=752.0638271$	OK	OK
$l=700$	$\beta=812$	$x=753.9230730$	NG	NG

3. DE（二重指数関数型積分法）の誤差評価

変数変換区間$[\varepsilon, 1-\varepsilon]$に対して区間幅$h$と分点数を定めると

一次元積分での誤差gは$g = \exp(-\dfrac{cN}{\ln(N)})$となる.ここで$g$は

絶対誤差なので,ここでは相対誤差と大きなオーダーの差

はないと仮定する。

またDE公式を与える変換$x = \phi(t)$では$\varepsilon = \phi(-t_0), 1-\varepsilon = \phi(t_0)$

となるt_0は$t_0 = \sinh^{-1}(\dfrac{1}{\pi}\ln(\dfrac{1}{\varepsilon}))$で与えられる。

$\varepsilon = 3.187\times 10^{-16}, 2^{-52} < \varepsilon < 2^{-51}$とすると$t_0 = 3.125$となる。

区間幅$h = 0.125$とすると分点数N_2は51, $h = 0.25$とすると分点数

$N_1 = 26$となる。また区間幅$h = 0.15625$とすると分点数N_2は41,

$h = 0.3125$とすると分点数$N_1 = 21$となる。

倍精度演算での評価ではこれらの区間幅,分点数で行うのは

妥当と言える。

通常解析解は不明なので,誤差評価式のcを求めるには2つの

分点数N_1, N_2での積分値y_1, y_2の差から求める必要がある。

$g = |y_1 - y_2| = \exp(-\dfrac{cN_1}{\ln(N_1)}) - \exp(-\dfrac{cN_2}{\ln(N_2)})$となるが,$N_1$と$N_2$の

大きさから$g = |y_1 - y_2| = \exp(-\dfrac{cN_1}{\ln(N_1)})$で$c$を求めても誤差は1%

前後である.これより$c = \dfrac{\ln(N_1)}{N_1}\ln(\dfrac{1}{g})$となる。

n次元の場合,分点数$(N_1, N_1, ..., N_1)$のときの積分値をy_1
分点数$(N_2, N_2, ..., N_2)$のときの積分値をy_2とする.
また分点数$(N_1, N_1, ..., N_1, N_2, N_1, ..., N_1)$　$(N_2 : i$番目の値$)$
のときの積分値をy_{1i}として誤差$g_i = | y_1 - y_{1i} |$とすると
$g \le \sqrt{\sum_{i=1}^{n} {g_i}^2} \le \sqrt{n} g_{\max}$　$(g_{\max} = \max_{1 \le i \le n} g_i)$で一次元のときの評価
式$g = | y_1 - y_2 | = \exp(-\frac{cN_1}{\ln(N_1)})$から$c$を求める事が出来る。
ここでg_iにばらつきがあればcの値は小さくなり,g_iにばらつきがなければ,cの値は大きくなるので,cの値からある次元の分点数を増やして精度を良くするという対策が取れる。
次ぎに端点付近の影響を見る場合,DD形式の4倍精度演算では$x = \phi(t)$のコーディングを一部修正するだけで変数変換区間$[\varepsilon, 1 - \varepsilon]$で$\varepsilon \to 2^{-1024}$とできるため(あくまで原理的にと言う意味で他の制限が加わる事もある),これまでと同じ方法で誤算の評価が出来る.
倍精度での変数変換区間$[\varepsilon_1, 1 - \varepsilon]$　$(0 < \varepsilon_1 << \varepsilon)$とすると
$t_1 = \sinh^{-1}(\frac{1}{\pi} \ln(\frac{1}{\varepsilon_1}))$で$t_1$を求める。ここで$\varepsilon_1 = \phi(-t_1)$.
DEの性質より,$\varepsilon_1 << \varepsilon$でも$\frac{t_1 - t_0}{h}$は小さくて済む。また
コーディングの手間を考え,hの値は変更せずに$-j + i$のjの値をt_1に対応する値に変更するだけに出来る。
例として,$h = 0.125$の場合,$N_2 = 51$から$N_2 = 61$にすると
$\varepsilon = 3.187 \times 10^{-16}$から,$\varepsilon_1 = 6.54 \times 10^{-55}$となる.ソースの修正部分は
$-26 + i => -36 + i$のみとなる.演算量はn次元積分で$(\frac{61}{51})^n$,
$n = 5$で約2.5倍の増加で済む。ここで変数変換区間$[\varepsilon, 1 - \varepsilon]$
と$[\varepsilon_1, 1 - \varepsilon]$での積分値により,0付近の影響をみる事ができる。

5次元積分laportad,laportafの誤算評価（ケース1）

laportadは2次元目と4次元目の誤差が大きい
laportafは誤差のばらつきはない。

DE誤差一覧表				
N=21 & 41	h=0.3125,h=0.15625			
			係数c	
誤差	laportad	laportaf	laportad	laportaf
gosa	8.715298E-06	2.189189E-07	1.69	2.22
gosa1	9.731227E-07	1.544159E-08		
gosa2	7.788459E-06	6.837236E-08		
gosa3	3.333854E-08	1.132839E-07		
gosa4	7.613123E-06	1.506979E-08		
gosa5	1.951379E-08	3.846251E-08		
sum gosa	1.642756E-05	2.506301E-07		
N=25 & N=49 h=0.3125,h=0.15625				
誤差	laportad	laportaf		
gosa	8.715298E-06	2.189189E-07		
gosa1	9.731227E-07	1.544159E-08		
gosa2	7.788459E-06	6.837236E-08		
gosa3	3.333854E-08	1.132839E-07		
gosa4	7.613123E-06	1.506979E-08		
gosa5	1.951379E-08	3.846251E-08		
sum gosa	1.642756E-05	2.506301E-07		

laportad.laportafともに0近傍の積分
への影響はない。

5次元積分laportad,laportafの誤算評価（ケース2）

laportadは2次元目と4次元目の誤差が大きい
laportafは誤差のばらつきはない。

DE誤差一覧表

N=26 & 51 h=0.25,h=0.125

			係数c	
誤差	laportad	laportaf	laportad	laportaf
gosa	1.179577E-06	1.001327E-08	1.71	2.31
gosa1	1.778147E-08	1.016066E-09		
gosa2	1.198222E-06	5.588154E-09		
gosa3	8.821643E-11	4.457539E-09		
gosa4	1.199953E-06	7.914422E-10		
gosa5	9.773612E-10	1.945273E-09		
sum gosa	2.417022E-06	1.379847E-08		

N=31 & N=61 h=0.25,h=0.125

誤差	laportad	laportaf
gosa	1.179577E-06	1.001327E-08
gosa1	1.778147E-08	1.016066E-09
gosa2	1.198222E-06	5.588154E-09
gosa3	8.821643E-11	4.457539E-09
gosa4	1.199953E-06	7.914422E-10
gosa5	9.773612E-10	1.945273E-09
sum gosa	2.417022E-06	1.379847E-08

laportad.laportafともに0近傍の積分
への影響はない。