Explanation of the answers, June 13

課題解答の解説

PROBLEM, June 13

- **Submit electronic file(s) via T2SCHOLAR until June 14**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

Choose one of the following PROBLEM 1 or PROBLEM 2

PROBLEM 1:

(i) Convert 101101_2 to base 10

(ii) Convert 3123_{10} to base 16

PROBLEM 2:

Choose one of the python programs given today (sum_error.py, sum.py, base.py).

・Explain what each block of the source code does,

or

・list up the source code parts that you cannot understand what they do or why they are needed.

今日配布したプログラム (sum_error-plt.py, sum.py, base.py) から1つを選び、以下のいずれかを答えよ

・ソースコードのそれぞれの部分が何をしているかを説明する

・ソースコードの中で理解できない部分、あるいは なぜそれが必要かわからない部分を述べよ

PROBLEM, June 13

- **Solve in this class time**
- **Submit electronic file(s) via T2SCHOLAR today after the class**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

Choose one of the following PROBLEM 1 or PROBLEM 2

PROBLEM 1:

(i) Convert 101101_2 to base 10

(ii) Convert 3123_{10} to base 16

PROBLEM 2:

Choose one of the python programs given today (sum_error.py, sum.py, base.py).

・Explain what each block of the source code does,

or

・list up the source code parts that you cannot understand what they do or why they are needed.

今日配布したプログラム (sum_error-plt.py, sum.py, base.py) から1つを選び、以下のいずれかを答えよ

・ソースコードのそれぞれの部分が何をしているかを説明する

・ソースコードの中で理解できない部分、あるいは なぜそれが必要かわからない部分を述べよ

PROBLEM, June 13

Choose one of the following PROBLEM 1 or PROBLEM 2

PROBLEM 1:

(i) Convert 110101_2 to base 10

> python base.py 110101 2 10

Convert 110101 in base 2 to base 10

0-th digit = 1_2:	+ 1*2^0 => +	1_10 =>	1_10
1-th digit = 0_2:	+ 0*2^1 => +	0_10 =>	1_10
2-th digit = 1_2:	+ 1*2^2 => +	4_10 =>	5_10
3-th digit = 0_2:	+ 0*2^3 => +	0_10 =>	5_10
4-th digit = 1_2:	+ 1*2^4 => +	16_10 =>	21_10
5-th digit = 1_2:	+ 1*2^5 => +	32_10 =>	53_10

PROBLEM, June 13

(ii) Convert 3053_{10} to base 16

> python base.py 3053 10 16

Convert 3053 in base 10 to base 10

0-th digit = 3_10: + 3^10 => + 3_10 => 3_10

1-th digit = 5_10: + 5^10 => + 50_10 => 53_10

2-th digit = 0_10: + 0^10 => + 0_10 => 53_10

3-th digit = 3_10: + 3^10 => + 3000_10 => 3053_10

Convert 3053 in base 10 to base 16

3053 = 190 * 16 + 13: base_16 => D

190 = 11 * 16 + 14: base_16 => ED

11 = 0 * 16 + 11: base_16 => BED

PROBLEM, June 13 #2

Choose one of the following PROBLEM 1 or PROBLEM 2

PROBLEM 1:

(i) Convert 101101_2 to base 10

> python base.py 101101 2 10

Convert 101101 in base 2 to base 10

0-th digit = 1_2:	+ $1*2^0$ => +	1_{10} =>	1_{10}
1-th digit = 0_2:	+ $0*2^1$ => +	0_{10} =>	1_{10}
2-th digit = 1_2:	+ $1*2^2$ => +	4_{10} =>	5_{10}
3-th digit = 1_2:	+ $1*2^3$ => +	8_{10} =>	13_{10}
4-th digit = 0_2:	+ $0*2^4$ => +	0_{10} =>	13_{10}
5-th digit = 1_2:	+ $1*2^5$ => +	32_{10} =>	45_{10}

PROBLEM, June 14 #2

(ii) Convert 3123_{10} to base 16

> python base.py 3123 10 16

Convert 3123 in base 10 to base 10

0-th digit = 3_10: + 3*10^0 => + 3_10 => 3_10

1-th digit = 2_10: + 2*10^1 => + 20_10 => 23_10

2-th digit = 1_10: + 1*10^2 => + 100_10 => 123_10

3-th digit = 3_10: + 3*10^3 => + 3000_10 => 3123_{10}

Convert 3123 in base 10 to base 16

3123 = 195 * 16 + 3: base_16 => 3

195 = 12 * 16 + 3: base_16 => 33

12 = 0 * 16 + 12: base_16 => C33

PROBLEM, June 13

Choose one of the following PROBLEM 1 or PROBLEM 2

PROBLEM 2:

Choose one of the python programs given today (sum_error.py, sum.py, base.py).

- Explain what each block of the source code does,

or

- list up the source code parts that you cannot understand what they do or why they are needed.

Python string format

python version 2.x <= No more support, should stop using ver 2 series

print “a=”, a Implemented as a built-in syntax

print(“a=”, a)

version 3.x print is implemented as a function. ‘()’ is mandatory.

print(“a=”, a)

Can use string format syntaxes:

1. “v=%3.1f i=%2d” % (v, i) C language type. Obsolete, not recommended
2. “v={:3.1f} i={:2d}.format(v, i) format method of string object
3. f“v={v:3.1f} i={i:2d} f-string literal. Version 3.6 or later
4. f”{v=:3.1f} {i=:2d} Version 3.9? or later too new to share the program with others

PROBLEM, June 13

base.py

```
# define the numerical equivalence between [0,35] and [0,Z]
base_chars = "0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ"
                        # list variable that converts an integer to a character
# define the numerical equivalence between [0,35] and [0,Z]
base_chars_dict = { # dict variable that converts a character to an integer
            "0": 0
            , “1": 1
-- cut –
            ,“Y": 34
            ,"Z": 35
   }

# Function that converts a character to an integer
def value_base(char, b):
            return base_chars_dict[char]
```

PROBLEM, June 13: Simpler program

base.py

```
# define the numerical equivalence between [0,35] and [0,Z]
base_chars = "0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ"
                    # list variable that converts an integer to a character

# Function that converts a character to an integer
def value_base(char, b):
        return list(base_chars).index(char)
```

PROBLEM, June 16

- **Submit electronic file(s) via T2SCHOLAR until June 14**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

(i) Calculate $dE(k)/dk$, $d^2E(k)/dk^2$, and effective mass m_e^*/m_0 from $E(k)$ in band.xlsx, and plot m_e^*/m_0 vs k.
Assume the lattice parameter is a = 4.0 Å.

(ii) Compare the results obtained by different h.

Effective mass

LCAO band

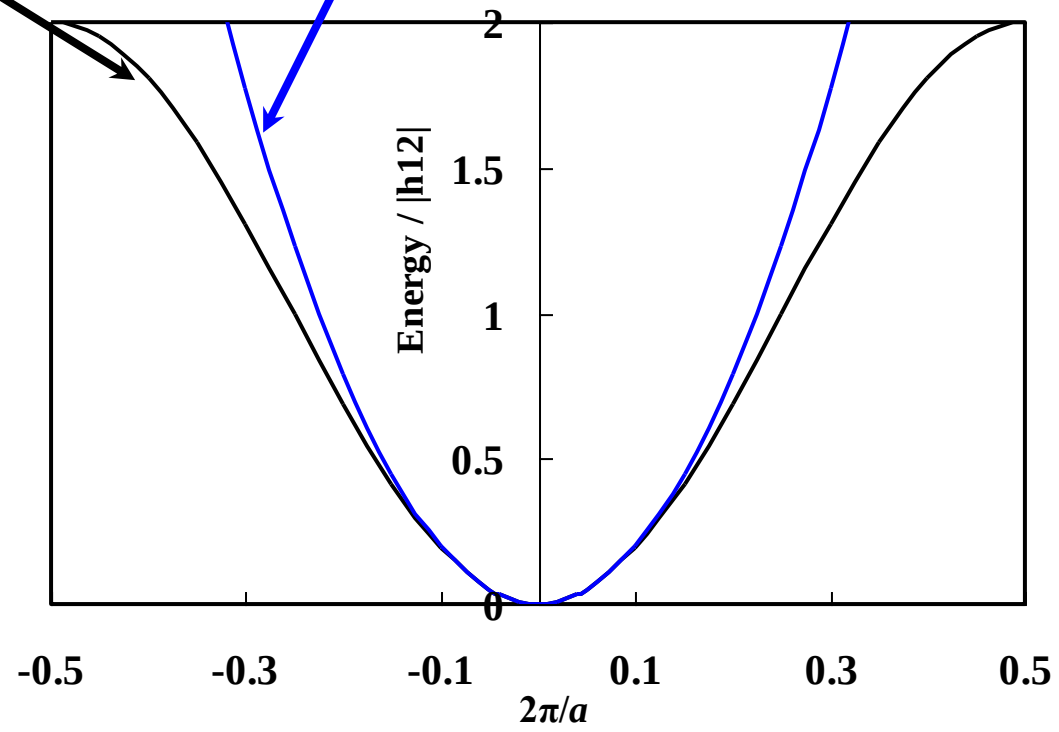
$$E(k) = \varepsilon_1 - 2|h_{12}|\cos(ka) \sim \varepsilon_1 - 2|h_{12}| + |h_{12}|a^2k^2 + O((ka)^4)$$

Free electron model

$$E(k) = E_0 + \frac{|\mathbf{P}|^2}{2m} = E_0 + \frac{\hbar^2}{2m}|\mathbf{k}|^2$$

$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{\partial^2 E_n(\mathbf{k})}{\partial k^2}$$

$$m^* = \frac{\hbar^2}{2|h_{12}|a^2}$$



Effective mass

k represents fractional coordinate in reciprocal unit cell:

generally expressed in the range **[-½ ½]**

Unit conversion $\boldsymbol{k_{real} = (2\pi/a)k}$

Note ***E*(*k*) is in eV**

$$m^* = \hbar^2 \left(\frac{\partial^2 E_J(\mathbf{k})}{\partial {k_{real}}^2} \right)^{-1} = \hbar^2 \left(\frac{2\pi}{a} \right)^2 \left(\frac{\partial^2 E_{eV}(\mathbf{k})}{\partial k^2} e \right)^{-1}$$

Very often effective mass is given by a ratio to the electron rest mass m_e^0 .

$$\boldsymbol{m^*/{m_e}^0 = \hbar^2 \left(\frac{\partial^2 E_J(k)}{\partial {k_{real}}^2} \right)^{-1} /{m_e}^0 = \hbar^2 \left(\frac{2\pi}{a} \right)^2 \left(\frac{\partial^2 E_{eV}(k)}{\partial k^2} e \right)^{-1} /{m_e}^0}$$

PROBLEM, June 16

- **Submit electronic file(s) via T2SCHOLAR until June 14**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

(i) Calculate $dE(k)/dk$, $d^2E(k)/dk^2$, and effective mass m_e^*/m_0 from $E(k)$ in band.xlsx, and plot m_e^*/m_0 vs k.
Assume the lattice parameter is a = 4.0 Å.

(ii) Compare the results obtained by different h.

See band_answer.xlsx

Effective mass

LCAO band

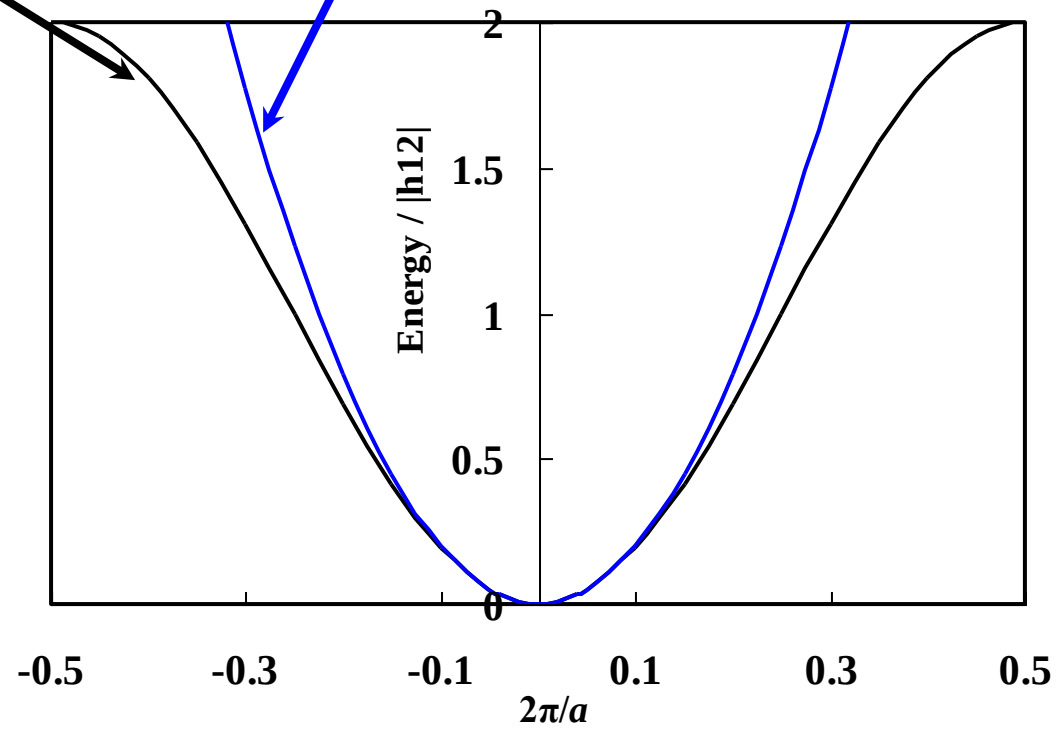
$$E(k) = \varepsilon_1 - 2|h_{12}|\cos(ka) \sim \varepsilon_1 - 2|h_{12}| + |h_{12}|a^2k^2 + O((ka)^4)$$

Free electron model

$$E(k) = E_0 + \frac{|\mathbf{P}|^2}{2m} = E_0 + \frac{\hbar^2}{2m}|\mathbf{k}|^2$$

$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{\partial^2 E_n(\mathbf{k})}{\partial k^2}$$

$$m^* = \frac{\hbar^2}{2|h_{12}|a^2}$$



Effective mass

k represents fractional coordinate in reciprocal unit cell:

generally expressed in the range **[-½ ½]**

Unit conversion $\boldsymbol{k_{real} = (2\pi/a)k}$

Note ***E*(*k*) is in eV**

$$m^* = \hbar^2 \left(\frac{\partial^2 E_J(\mathbf{k})}{\partial {k_{real}}^2} \right)^{-1} = \hbar^2 \left(\frac{2\pi}{a} \right)^2 \left(\frac{\partial^2 E_{eV}(\mathbf{k})}{\partial k^2} e \right)^{-1}$$

Very often effective mass is given by a ratio to the electron rest mass m_e^0 .

$$\boldsymbol{m^*/{m_e}^0 = \hbar^2 \left(\frac{\partial^2 E_J(k)}{\partial {k_{real}}^2} \right)^{-1} /{m_e}^0 = \hbar^2 \left(\frac{2\pi}{a} \right)^2 \left(\frac{\partial^2 E_{eV}(k)}{\partial k^2} e \right)^{-1} /{m_e}^0}$$

Numerical differentiation: Accuracy

$$\frac{df(x)}{dx} \sim \frac{f(x+h)-f(x)}{h}$$

Error: $$\frac{f(x+h)-f(x)}{h} = \frac{df(x)}{dx} + \frac{1}{2}\frac{d^2 f(x)}{dx^2}h + \frac{1}{3!}\frac{d^3 f(x)}{dx^3}h^2 + O(h^4)$$

$$\frac{df(x)}{dx} \sim \left[\frac{f(x+h)-f(x)}{h} + \frac{f(x)-f(x-h)}{h}\right]/2 = \frac{f(x+h)-f(x-h)}{2h}$$

$$f(x+h) = f(x) + \frac{df(x)}{dx}h + \frac{1}{2}\frac{d^2 f(x)}{dx^2}h^2 + \frac{1}{3!}\frac{d^3 f(x)}{dx^3}h^3 + O(h^4)$$

$$f(x-h) = f(x) - \frac{df(x)}{dx}h + \frac{1}{2}\frac{d^2 f(x)}{dx^2}h^2 - \frac{1}{3!}\frac{d^3 f(x)}{dx^3}h^3 + O(h^4)$$

Error: $$\frac{f(x+h)-f(x-h)}{2h} = \frac{df(x)}{dx} + \frac{1}{3!}\frac{d^3 f(x)}{dx^3}h^2 + O(h^3)$$

Second differential (二階微分)

If calculate 2nd differential using forward differences with the 1st and the 2nd differentials … (一階微分を前身差分で計算してから二階微分を前進差分で計算すると･･･)

$$\frac{d^2x(t)}{dt^2} = \frac{\frac{dx}{dt}(t+\Delta t) - \frac{dx}{dt}(t)}{\Delta t} \sim \frac{\frac{x(t+2\Delta t)-x(t+\Delta t)}{\Delta t} - \frac{x(t+\Delta t)-x(t)}{\Delta t}}{\Delta t} = \frac{x(t+2\Delta t)-2x(t+\Delta t)+x(t)}{\Delta t^2}$$

Use central difference

$$\frac{d^2x(t)}{dt^2} = \frac{\frac{dx}{dt}(t+\Delta t/2) - \frac{dx}{dt}(t-\Delta t/2)}{\Delta t} \sim \frac{\frac{x(t+\Delta t)-x(t)}{\Delta t} - \frac{x(t)-x(t-\Delta t)}{\Delta t}}{\Delta t} = \frac{x(t+\Delta t)-2x(t)+x(t-\Delta t)}{\Delta t^2}$$

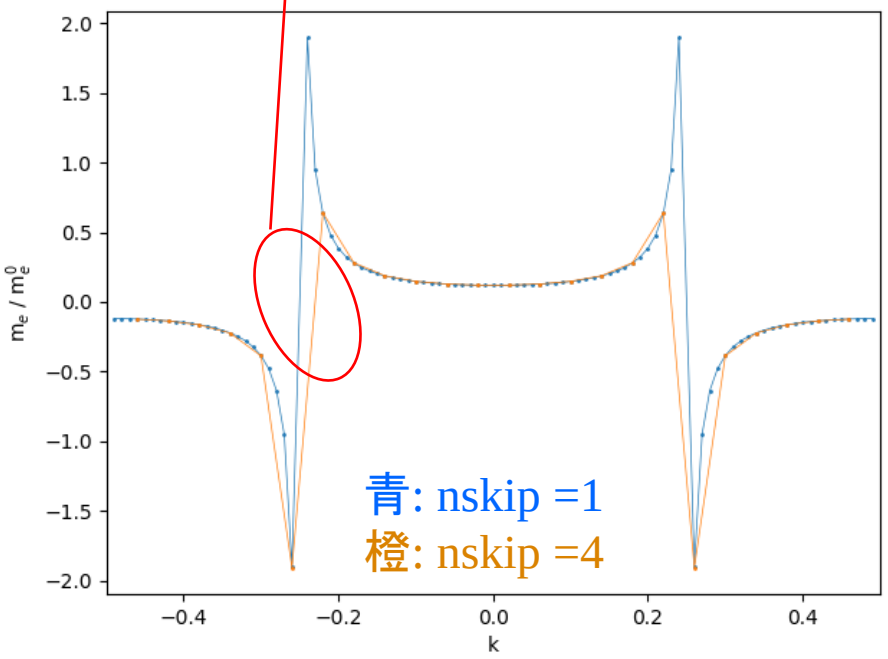
Note: These two formula are offset in *t* by *Δt*

2つの式では、横軸が*Δt*ずれるので注意！

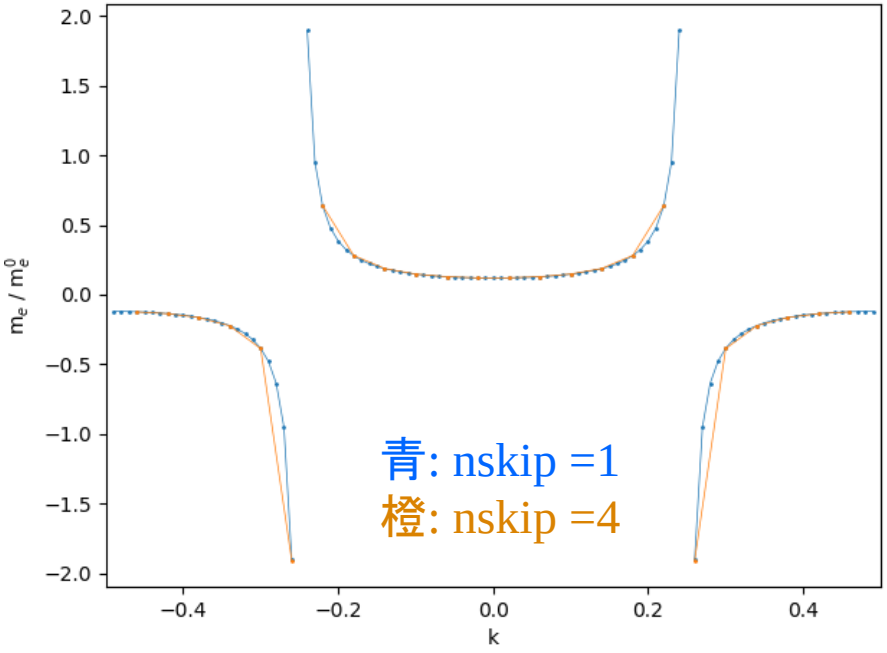
プログラム (抜粋)

python EffectiveMass.py

These lines should not be drawn
描いてはいけない線



Data points can be disconnected by inserting None values
データに None (未定義値) を挿入することで
描いてはいけない線を消した



プログラム (抜粋)

EffectiveMass.py

```
#==================================
# parameters
#==================================
#a   = 4.0 # A
a = 4.0e-10 #m

infile = 'band.csv'
#有効質量の符号が変わる点をつなぐかどうかのフラグ
cutline = 1

def read_csv(fname):
    x = []
    y = []
    with open(fname) as f:
        fin = csv.reader(f)
        xlabel, ylabel, = next(fin)
        for row in fin:
            try:
                x.append(float(row[0]))
                y.append(float(row[1]))
            except:
                print("Warning: Invalid float data [{}] or [{}]".format(row[0],
row[1]))

    return xlabel, ylabel, x, y

def main():
    klabel, Elabel, k, E = read_csv(infile)
#入力データで使う変数を計算
    nk = len(k)
    dk = k[1] - k[0]
#共通の定数は先に計算
    km = hbar * hbar * (pi2 / a)**2.0
#微分の精度を比較するため、h = nskip*dk にする
    nskip = 1
    xk  = []
    ymc = []
#符号の変化を検出するため、符号変数を用意
signprev = None
for i in range(nskip, nk - nskip, nskip):
#2階微分を計算
        d2Edk2c = (E[i+nskip] + E[i-nskip] - 2 * E[i]) * e / pow(nskip *
dk, 2.0)
#2回微分はゼロになることがあるので、まずは1/m*を計算
        minv = d2Edk2c / km
        print(i, E[i-1], E[i], E[i+1], minv)
#1/m*が1/meより非常に小さければ、m*は計算しない
        if abs(minv) <= 1.0e20:  # << 1.0/me ~ 1e30
#符号が反転する場所でグラフの線を切断するときは
#Noneデータを追加する。
            if cutline:
                xk.append(k[i])
                ymc.append(None)
#反転した符号を記録
                signprev = -signprev
                continue
        else:
            m = km / d2Edk2c

        if signprev is None:
            signprev = m
        elif signprev * m < 0.0:
            if cutline:
```

プログラム (抜粋)

EffectiveMass.py

```
#共通の定数は先に計算
   km = hbar * hbar * (pi2 / a)**2.0
#微分の精度を比較するため、h = nskip*dk にする
    nskip = 1
    xk  = []
    ymc = []
#符号の変化を検出するため、符号変数を用意
signprev = None
for i in range(nskip, nk - nskip, nskip):
#2階微分を計算
        d2Edk2c = (E[i+nskip] + E[i-nskip] - 2 * E[i]) * e / pow(nskip *
dk, 2.0)
#2回微分はゼロになることがあるので、まずは1/m*を計算
        minv = d2Edk2c / km
        print(i, E[i-1], E[i], E[i+1], minv)
#1/m*が1/meより非常に小さければ、m*は計算しない
        if abs(minv) <= 1.0e20:  # << 1.0/me ~ 1e30
#符号が反転する場所でグラフの線を切断するときは
#Noneデータを追加する。
            if cutline:
                xk.append(k[i])
                ymc.append(None)
#反転した符号を記録
                signprev = -signprev
                continue
        else:
            m = km / d2Edk2c

#符号が反転する場所でグラフの線を切断するときは
#Noneデータを追加する。
        if signprev is None:
#signprevが 初期値 None である場合は 符号の最初の値を代入
            signprev = m
        elif signprev * m < 0.0:
            if cutline:
                xk.append(k[i])
                ymc.append(None)
#反転した符号を記録
            signprev = m

        xk.append(k[i])
        ymc.append(m / me)

    plt.plot(xk,  ymc,  linewidth = 0.5, marker = 'o', markersize = 1.0,
label = 'nskip = 1')
    plt.xlabel(klabel)
    plt.ylabel("m$_e$ / m$_e^0$")
    plt.xlim([-0.5, 0.5])
#    plt.ylim([-0.5, 0.5])
    plt.tight_layout()

    plt.pause(0.1)
    print("Press ENTER to exit>>", end = '')
    input()


if __name__ == "__main__":
    main()
```

PROBLEM, June 20

- **Submit electronic file(s) via T2SCHOLAR until June 22**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

(i) By filling the dx/dt and the $x(t)$ columns in diffeq.xlsx, solve $dx(t) / dt = -x(t)\sin(\pi t)$ using the Euler method.

Conditions:

t starts from 0 and ends at 3.0 with the time step of 0.1.

$x(0) = 1.0$

PROBLEM, June 20

PROBLEM:

(i) By filling the *dx*/*dt* and the *x*(*t*) columns in diffeq.xlsx, solve $dx(t)/dt = -x(t)\sin(\pi t)$ using the Euler method.

Conditions:

***t* starts from 0 and ends at 3.0 with the time step of 0.1.**

***x*(0) = 1.0**

Euler formula: $\frac{dx(t)}{dt} = f(x(t), t) = -x(t)\sin(\pi t)$

$$x(t + \Delta t) = x(t) + \Delta t \cdot f(\boldsymbol{x(t), t})$$

Typical mistake:

$$x(t + \Delta t) = x(t) + \Delta t \cdot f(x(\boldsymbol{t + \Delta t}), \boldsymbol{t + \Delta t})$$

This calculation is not possible if $f(x, t)$ includes *x* explicitly.

See diffeq2_answer.xlsx

PROBLEM, June 23

- **Submit electronic file(s) via T2SCHOLAR in 2 days**
  (If T2SCHOLAR doesn't work, send the files to tkamiya@msl.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

Smoothen the data DOS(E) in dos.xlsx
by simple average method and polynomial fit method.

Add them and plot the raw DOS(E) and the smoothed data in an Excel file.

You can choose smoothing parameters as you like, but explicitly descrive them.

Submit the excel file.

PROBLEM, June 23

PROBLEM:

Smoothen the data DOS(E) in dos.xlsx
by simple average method and polynomial fit method.

Add them and plot the raw DOS(E) and the smoothed data in an Excel file.

You can choose smoothing parameters as you like, but explicitly descrive them.

Submit the excel file.

See dos_smoothing_answer.xlsx

Smoothing

Simple moving average (2m+1 points)

$$y_{i,smoothed} = \frac{1}{2m+1} \sum_{j=i-m}^{i+m} y_j$$

Weighted moving average (2m+1 points)

$$y_{i,smoothed} = \frac{1}{\sum_{j=i-m}^{i+m} w_i} \sum_{j=i-m}^{i+m} w_i y_j$$

Order 2 and 3 polynomial fit using (2*m*+1) points

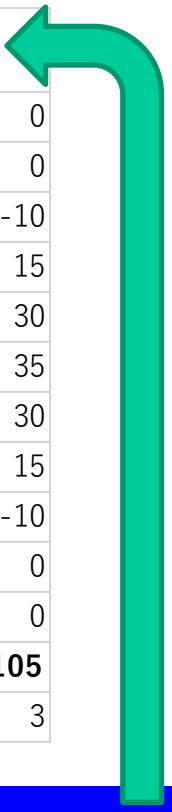
$w_{23}(j) = 3m(m+1) - 1 - 5j^2$ $\quad j$ = -m, ⋯, -1, 0, 1, ⋯, m

$W_{23} = (4m^2 - 1)(2m + 3)/3$

$$y_{i,smoothed} = \frac{1}{W_{23}} \sum_{j=i-m}^{i+m} w_{23}(j) y_j$$

Weights

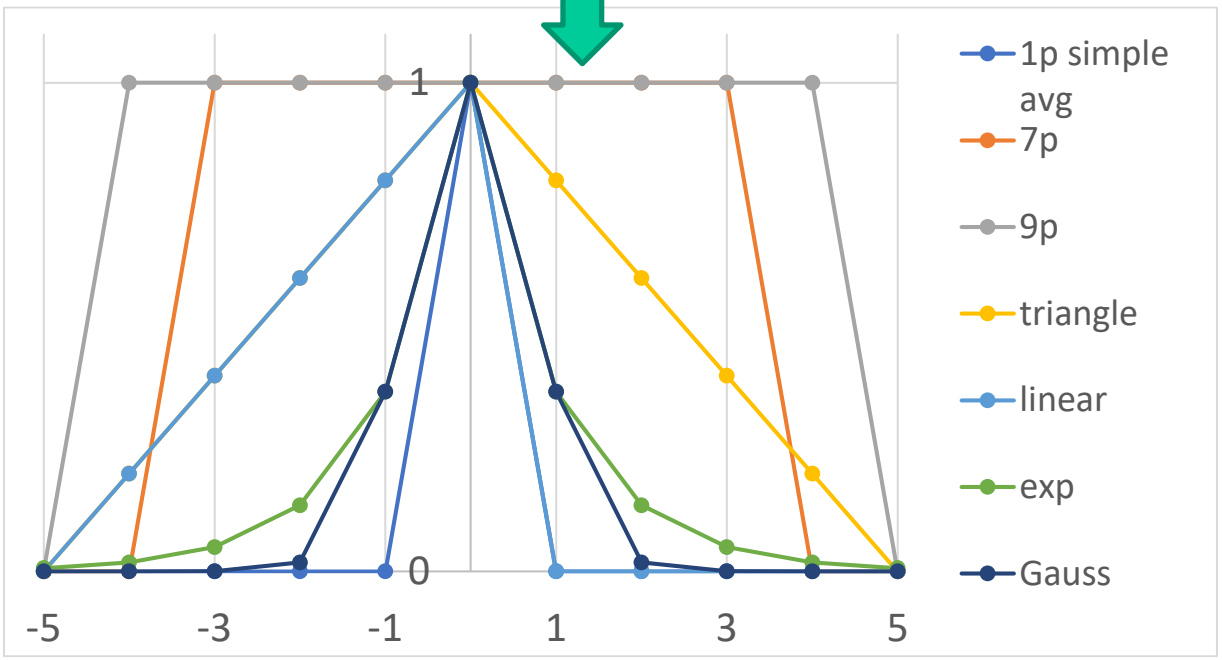
i-i0	1p simple avg	7p	9p	triangle	linear	exp	Gauss	3p order 3 polynomial	5p	7p
-5	0	0	0	0	0	0.006738	1.39E-11	0	0	0
-4	0	0	1	0.2	0.2	0.018316	1.13E-07	0	0	0
-3	0	1	1	0.4	0.4	0.049787	0.000123	0	0	-10
-2	0	1	1	0.6	0.6	0.135335	0.018316	0	-3	15
-1	0	1	1	0.8	0.8	0.367879	0.367879	0	12	30
0	1	1	1	1	1	1	1	5	17	35
1	0	1	1	0.8	0	0.367879	0.367879	0	12	30
2	0	1	1	0.6	0	0.135335	0.018316	0	-3	15
3	0	1	1	0.4	0	0.049787	0.000123	0	0	-10
4	0	0	1	0.2	0	0.018316	1.13E-07	0	0	0
5	0	0	0	0	0	0.006738	1.39E-11	0	0	0
W	**1**	**7**	**9**	**5**	**3**	**2.15611**	**1.77264**	**5**	**35**	**105**
							m=	1	2	3



Equivalent to convolution using these functions

Weights

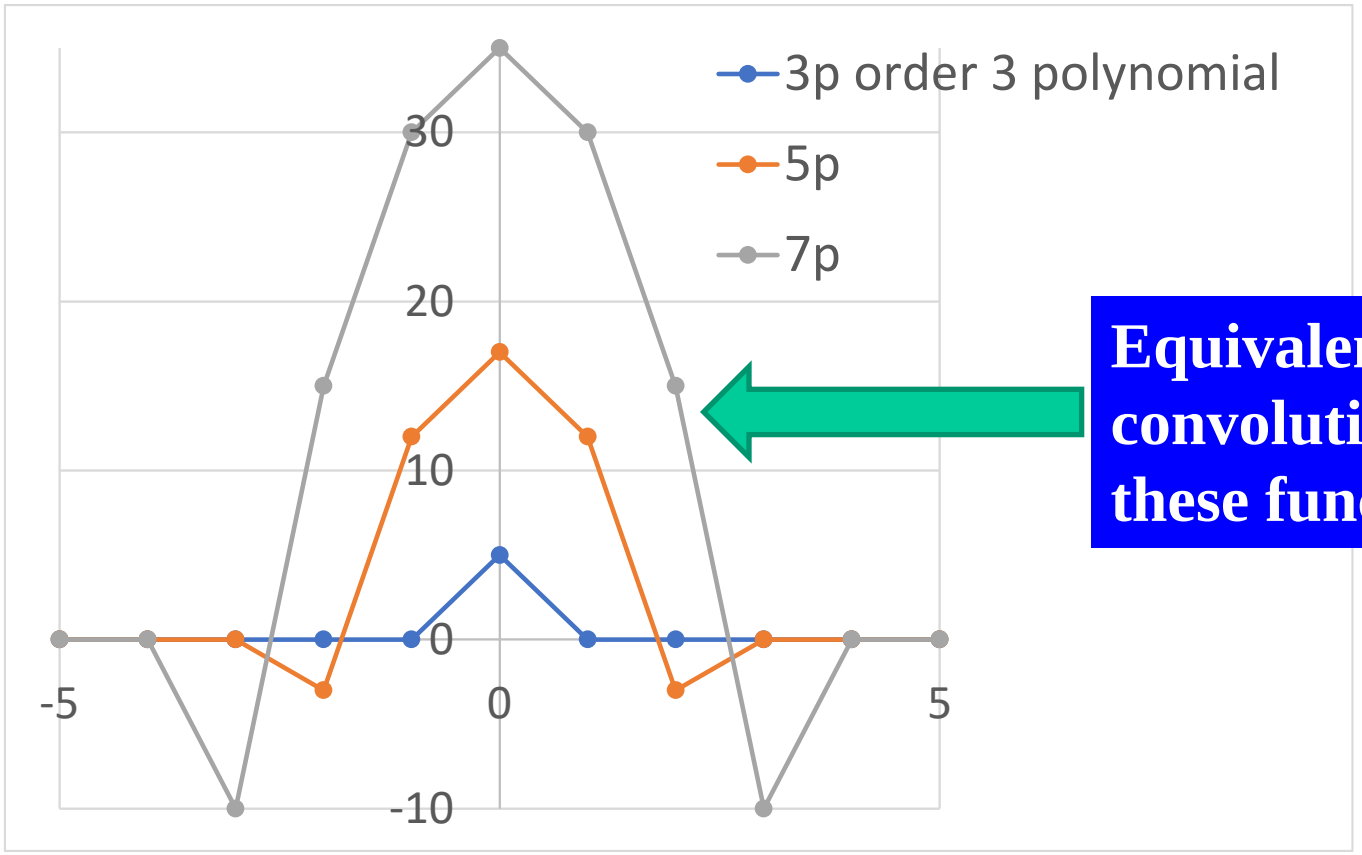
Order 2 and 3 polynomial fit using (2*m*+1) points



Equivalent to convolution using these functions

Weights

Order 2 and 3 polynomial fit using (2*m*+1) points



Equivalent to convolution using these functions

PROBLEM, June 27

- **Submit electronic file(s) via T2SCHOLAR in 2 days**
  (If T2SCHOLAR doesn't work, send the files to kamiya.t.aa@m.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

Solve $5\cos(x) - x = 0$.

(i) Plot the functions $y = 5\cos(x)$ and $y = x$ in the range $x = 0 - 3$, find an initial x for Newton-Raphson method.

(ii) Solve $5\cos(x) - x = 0$ by Newton-Raphshon method at least with four significant digits.

PROBLEM, June 27

- **Submit electronic file(s) via T2SCHOLAR in 2 days**
  (If T2SCHOLAR doesn't work, send the files to kamiya.t.aa@m.titech.ac.jp. In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM:

Solve $5\cos(x) - x = 0$.

(i) Plot the functions $y = 5\cos(x)$ and $y = x$ in the range $x = 0 - 3$, find an initial x for Newton-Raphson method.

(ii) Solve $5\cos(x) - x = 0$ by Newton-Raphshon method at least with four significant digits.

Newton-Raphson method:

$f(x_0+dx) = f(x_0) + dx\, f'(x_0) \sim 0$

$\Rightarrow \quad x_1 = x_0 + dx = x_0 - f(x_0) / f'(x_0)$

See equation_answer.xlsx

PROBLEM, June 30

- **Submit electronic file(s) via T2SCHOLAR in 2 days**
  (If T2SCHOLAR doesn't work, send the files to kamiya.t.aa@m.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM: Answer can be in Japanese or English

(i) Which algorism you can apply to your research? Describe in a few lines

(ii) Optional: Propose if you have any other numerical analysis you want to learn in Computational Materials Science

(iii) Optional: Propose if you have any python program (should be simple) you want to learn

PROBLEM, June 30

- **Submit electronic file(s) via T2SCHOLAR in 2 days**
  (If T2SCHOLAR doesn't work, send the files to kamiya.t.aa@m.titech.ac.jp.
  In this case, file name must include your STUDENT ID and FULL NAME)

PROBLEM: Answer can be in Japanese or English

(i) Which algorism you can apply to your research? Describe in a few lines

(ii) Optional: Propose if you have any other numerical analysis you want to learn in Computational Materials Science

(iii) Optional: Propose if you have any python program (should be simple) you want to learn

tkProg web page

http://conf.msl.titech.ac.jp/D2MatE/D2MatE_programs.html

- **Python 3.9+**
- **GUI interface for command-line programs**
- **Optimization, Regression (Fitting) (Least-squares method, Machine-learning regression)**
  Smoothing, Fourier transformation etc

- Top page
- 共通ファイル・単位など
  - 神谷・片瀬研共通単位表
- 標準ディレクトリ構成
- Launcherプログラミング
- D2MatE拠点限定ページ
- 更新履歴

公開プログラム

ラウンチャプログラム Launcher.py から 各プログラムを実行できるように パッケージを配布しています

- **パッケージ記号**
  - **A**: 一般(all)
  - **C**: 神谷・片瀬研
  - **D**: D^{2}MatE
- **その他記号**
  - **pl**: perlが必要
  - **L**: Linuxのみ動作保証

注: インストールトラブル、エラーやバグを見つけた場合は、このページに従って報告をお願いします。

- **インストール方法**

1. python
2. pythonモジュール
3. PHYSBO
4. tkProg
5. 動作確認

智慧とデータが拓くエレクトロニクス新材料開発拠点　公開・非公開プログラム情報
(Data Driven Materials Research Institute for Electronics)

質問、要望、バグ報告などの連絡先: 神谷 利夫 tkamiya@msl.tech.ac.jp
東京工業大学 国際先駆研究センター 元素戦略MDX研究センター 教授

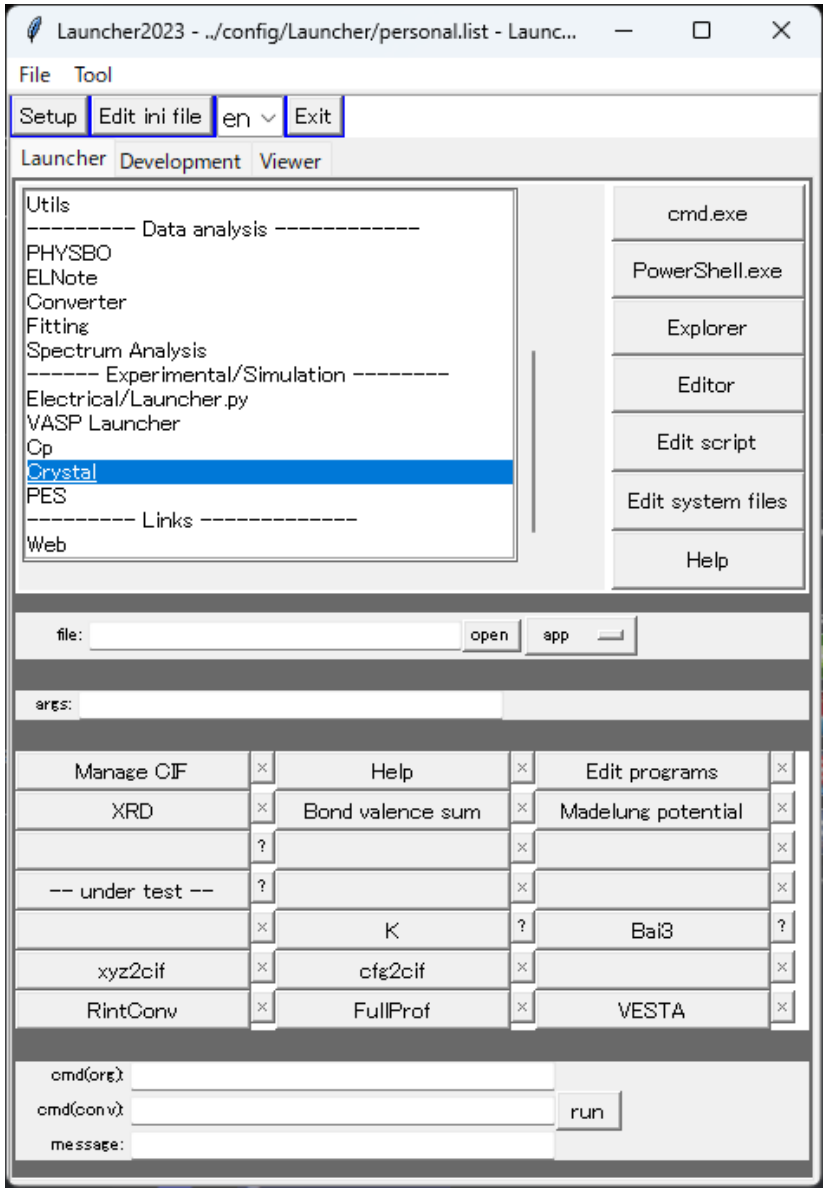
News

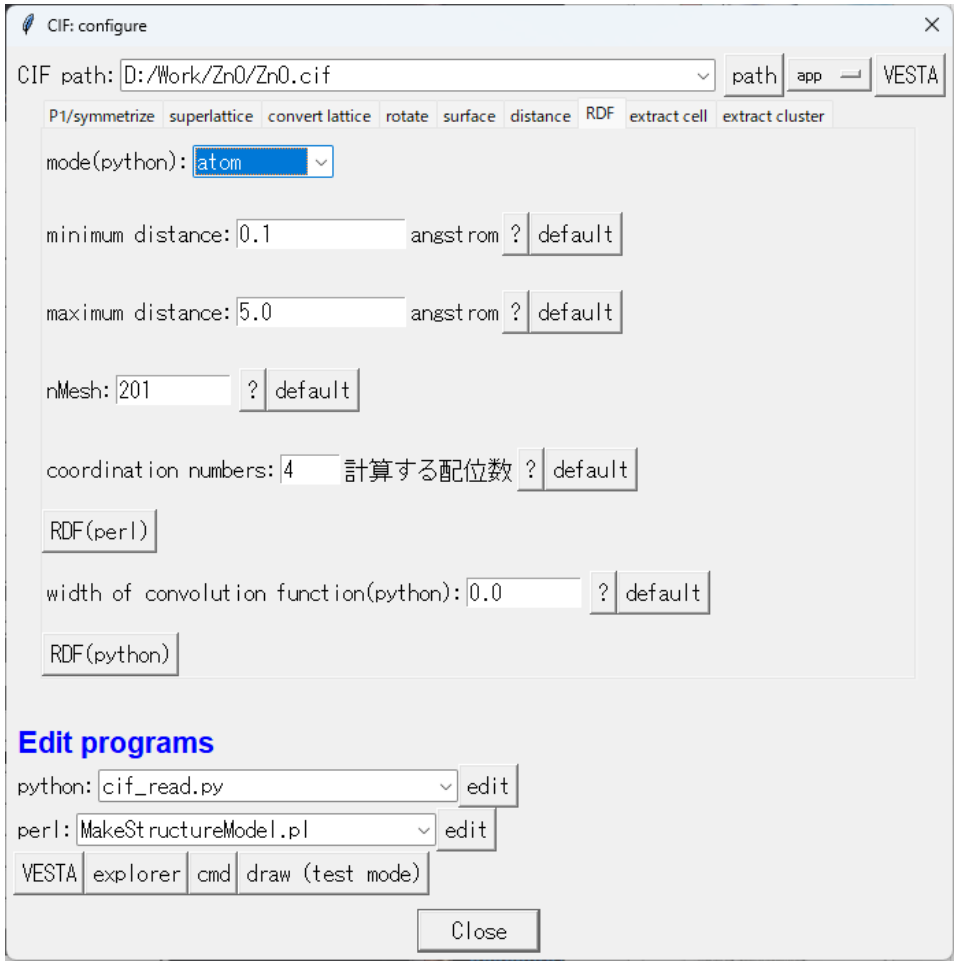
- **New!** 2023/6/29 13:57 一般向けパッケージを更新しました。以下の機能が追加されています。
  - Linux対応 (xterm が必要。一部プログラムでは perl およびライブラリが必要)
  - Crystal (CIFファイル) 関連プログラム (一部プログラムでは perl およびライブラリが必要)
  - VASP計算・後処理支援プログラム (一部プログラムでは perl およびライブラリが必要)
  - 電気物性関連プログラム
  - 比熱の温度依存性

  一部のプログラムは perlのインストールが必要になります。
  Linux版のインストールは、左メニューの 「tkProg Linux版インストール」 をご参照ください。

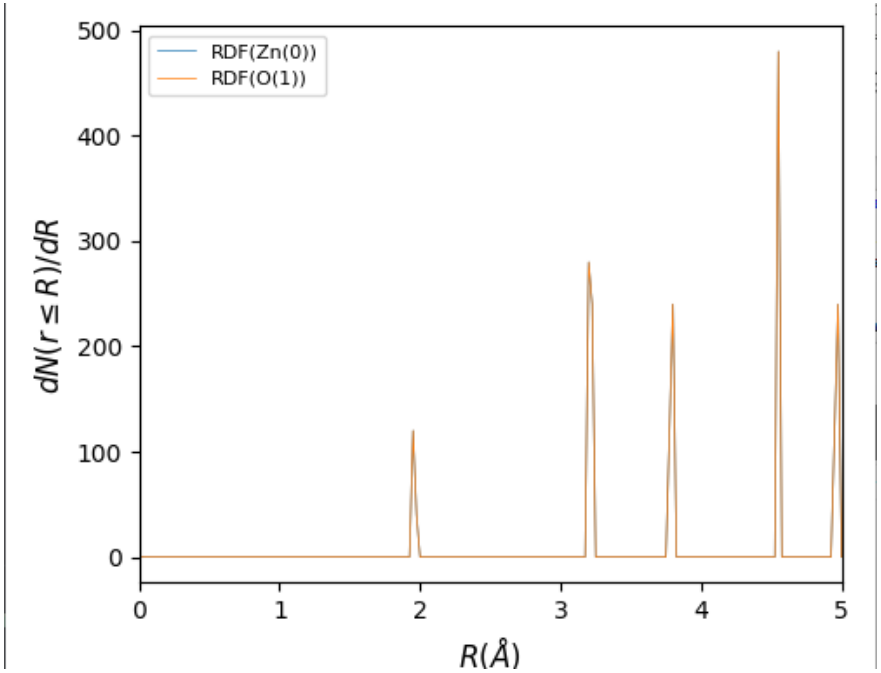
- **New!** 2023/6/29 14:20 6/29 11:13公開のパッケージではVASP支援機能が読み込めなかったので、修正版に差し替え。
  New! 2023/6/29 11:13 D^{2}MatE限定パッケージを D2MatE メンバー限定ページ で公開しました。以下の機能が追加されています。
  - Linux対応 (xterm が必要。一部プログラムでは perl およびライブラリが必要)
  - Crystal (CIFファイル) 関連プログラム (一部プログラムでは perl およびライブラリが必要)
  - VASP計算・後処理支援プログラム (一部プログラムでは perl およびライブラリが必要)
  - 電気物性関連プログラム
  - 比熱の温度依存性

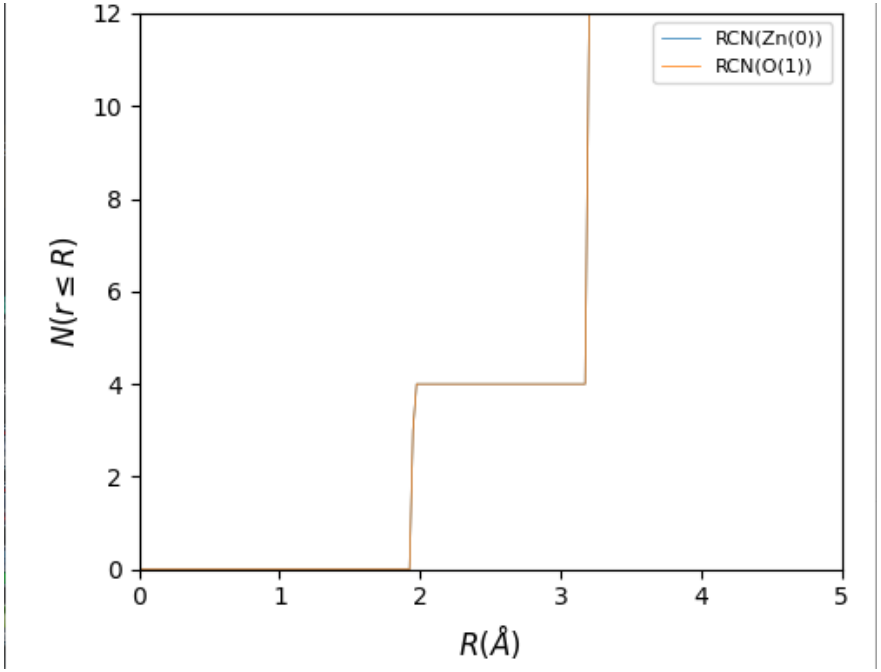
Q: MDの分析手法: RDF?





A: MDの分析手法: RDF

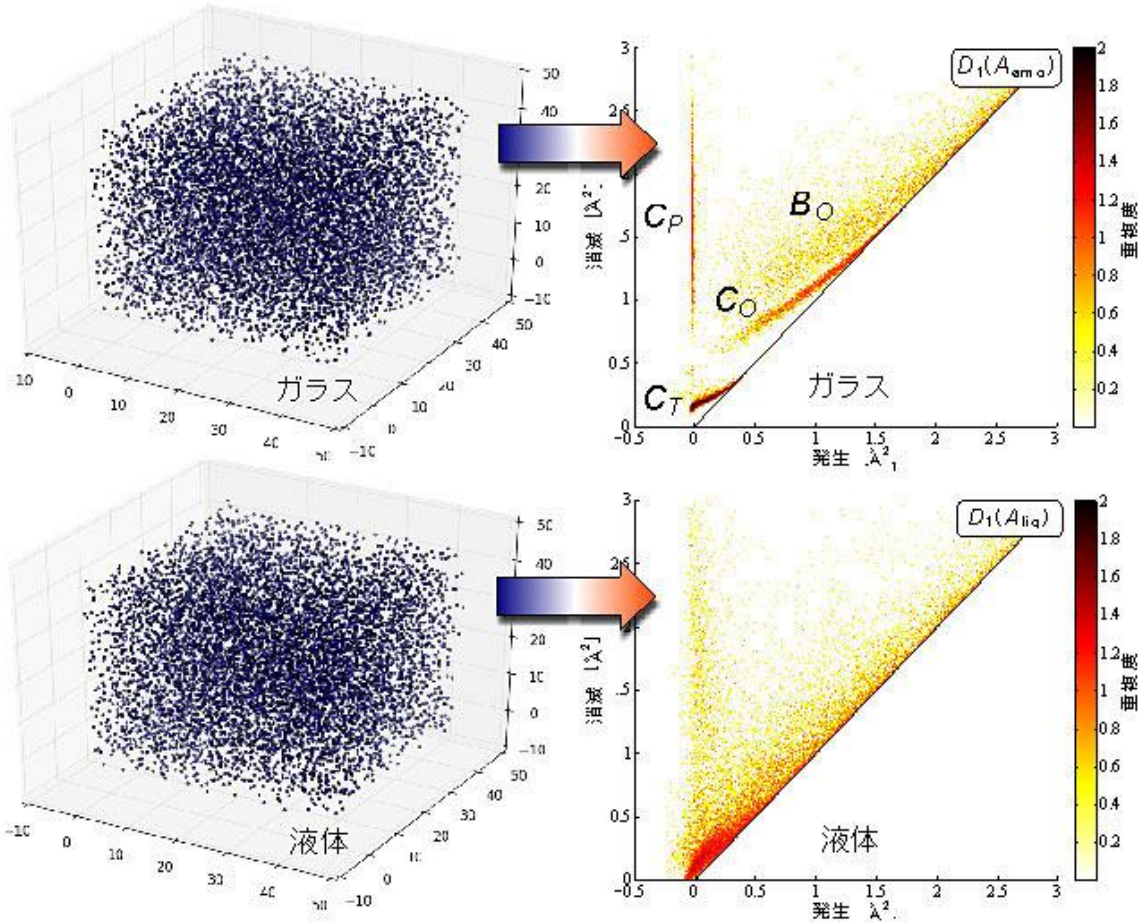




A: MDの分析手法: Persistent homology

Developed to characterize amorphous structures

https://www.jst.go.jp/pr/announce/20160614/index.html



Q: Python GUI programming

- **tkinter: Tk interface of python**
  - **Tk: The basic GUI APIs for X Window**
  - **Many languages support**
  - **Newer GUI APIs employ similar interfaces**
  - **Standard of python, but bit old fashioned outlook**
    (themed Tk tkinter.ttk improves)
- **PyGTK: GTK interface of python**
  - **GTK: GNU-developing GUI APIs**
    (formerly The GIMP Toolkit)
  - **Smart outlook (themed)**
- **PyQt: Qt interface of python**
  - **Qt: developed by a company Qt**
    Free version available, but be careful for license
  - **Many libraries**

Q: Python GUI programming: PySimpleGUI

Just try PySimpleGUI

- **https://fuji-pocketbook.net/pysimplegui/**
- **https://qiita.com/ku_a_i/items/20004800adc78fa94a0b**

Q: Python GUI programming: tkinter

- **http://conf.msl.titech.ac.jp/Lecture/python/tutorial-optimize/index-python-optimize-gui.html**

pythonによる最小二乗法・最適化問題 ＧＵＩプログラミング

関連Web

1. 計算材料科学特論 2022

全ファイルのZIPアーカイブ: tutorial-optimize.zip

pythonの基本

1. トップページ
2. pythonのインストール方法
   Install python
3. Pythonの基本変数型
4. pythonの起動と対話モード

pythonによる最適化GUIプログラミング

a) 基本

1. CSVファイルの読み込みとグラフプロット
   01-readcsv.py
   02-plotcsv.py

pythonによる最小二乗法・最適化問題 ＧＵＩプログラミング

ファイル

- 全ファイルのZIPアーカイブ: tutorial-optimize.zip

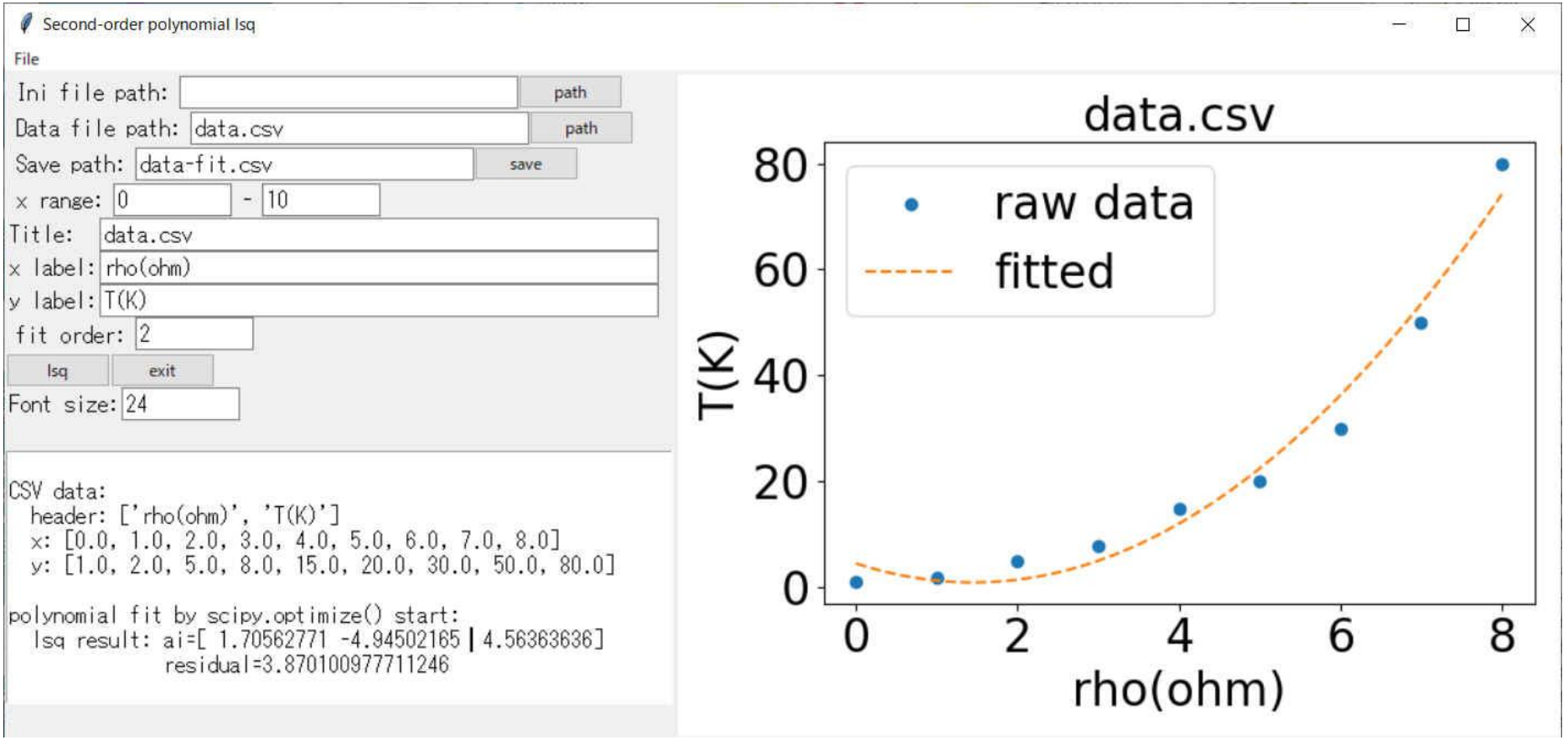
1. 最小二乗法講義スライド: 講義-プログラム-数値計算2019E.pptx

pythonプログラム

a) 基本

1. 入力データ (CSVファイル)
   data.csv
2. CSVの読み込み
   01-readcsv.py
3. CSVから読み込んだデータをグラフ(散布図)にプロット
   02-plotcsv.py

Q: Python GUI programming: tkinter + matplotlib



Q: Machine learning

Machine learning

- **Bayesian optimization**
  http://conf.msl.titech.ac.jp/D2MatE/bayes_gp/bayes_gp.html
- **Machine learning regressions**
  http://conf.msl.titech.ac.jp/D2MatE/ML_regression/ML_regression.html

Materials informatics

It covers a wide range of data analysis for materials science and engineering.

I cannot cover it in the CMS101.

A: Text books
Machine learning, data processing

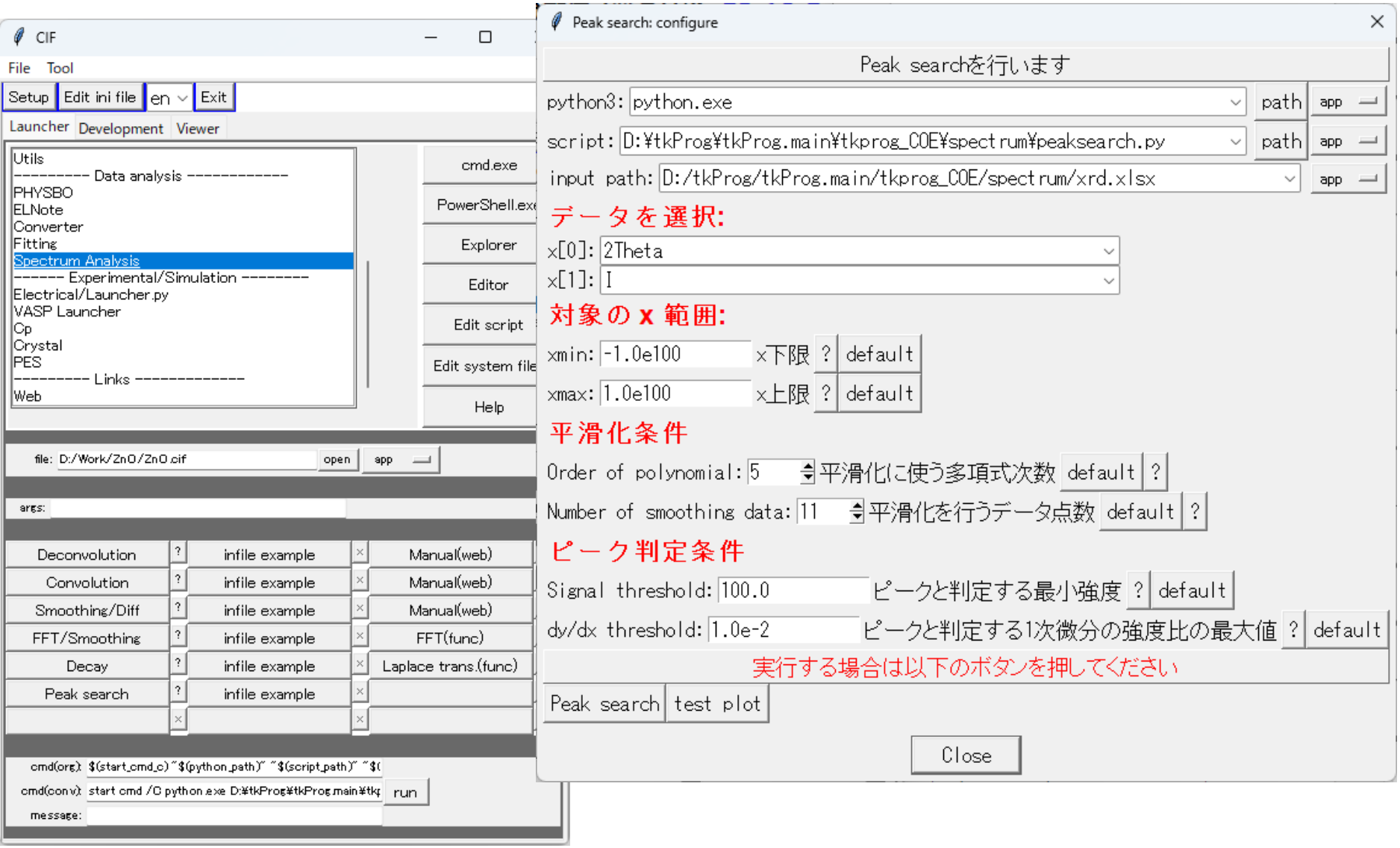
下山輝昌

株式会社Iroribi 代表取締役。日本電気株式会社(NEC)の中央研究所にてハードウェアの研究開発に従事した後、独立。機械学習を活用したデータ分析やダッシュボードデザイン等に裾野を広げ、データ分析コンサルタントとして幅広く案件に携わる。それと同時に、最先端にはテクノロジーとビernet of Things(IoT)、

Graduated Tokyo Tech, the former materials science course, now operating a data analysis company

Q: Peak search program

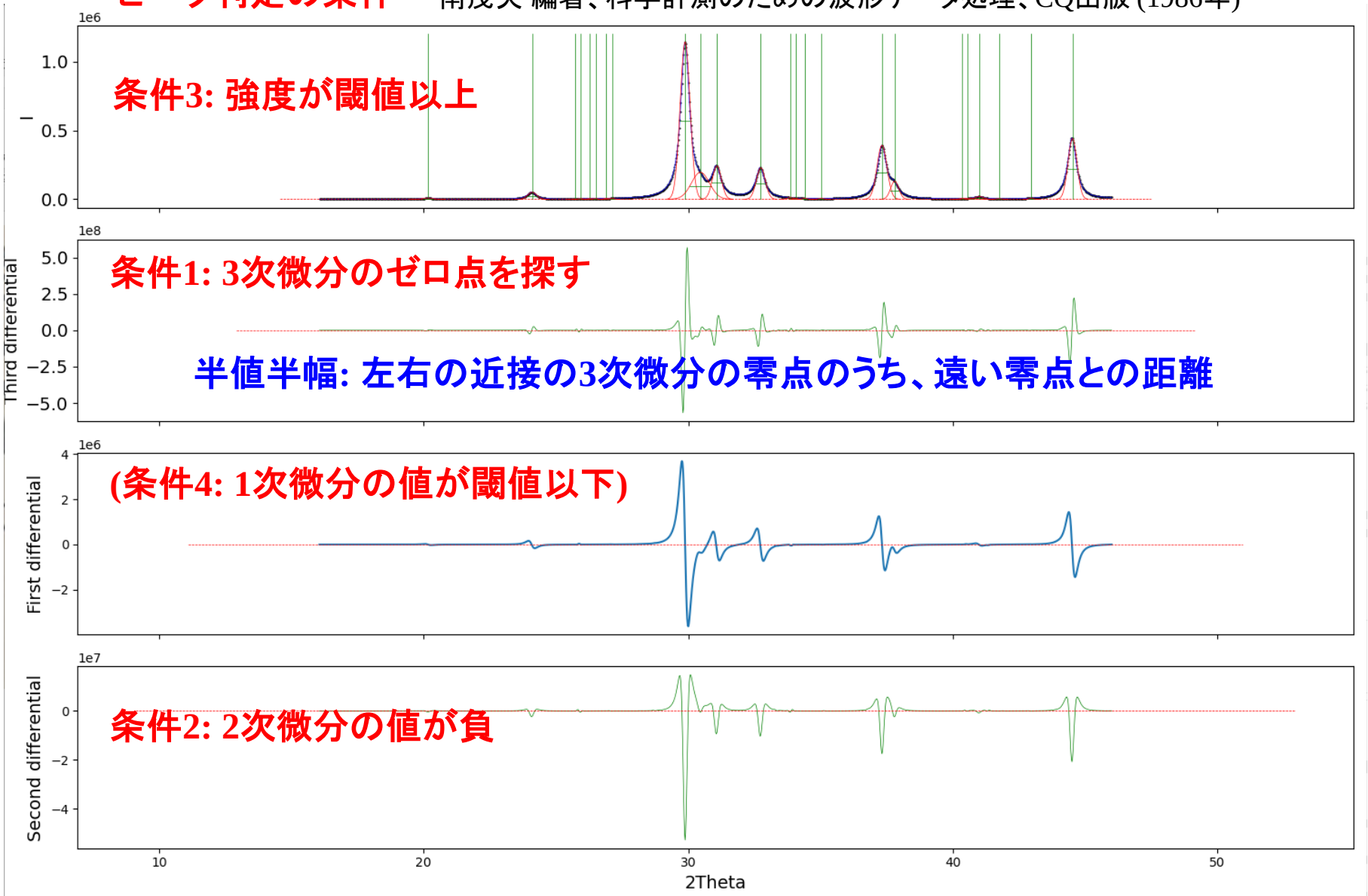
- **http://conf.msl.titech.ac.jp/D2MatE/PeakSearch/PeakSearch.html**



A: Peak search program

[tkProg]¥tkprog_base¥spectrum¥peak_search. py

ピーク判定の条件　南茂夫 編著、科学計測のための波形データ処理、CQ出版 (1986年)



Q: Methfessel-Paxton and tetrahedron method

These are used to integrate / smear E(k) obtained by band calculations

purposes of smearing used in band calculations.

1. **Increasing the accuracy of 1st BZ integration (interpolation)**
   Tetrahedron method

2. **Stabilize convergence of SCF (distribution)**
   Gauss smearing, Fermi smearing

3. **Make the DOS display easier to read (smoothing)**
   Polynomial fitting may help
   but convolution (smearing) and tetrahedron method are commonly used

Problem of smearing by convolution

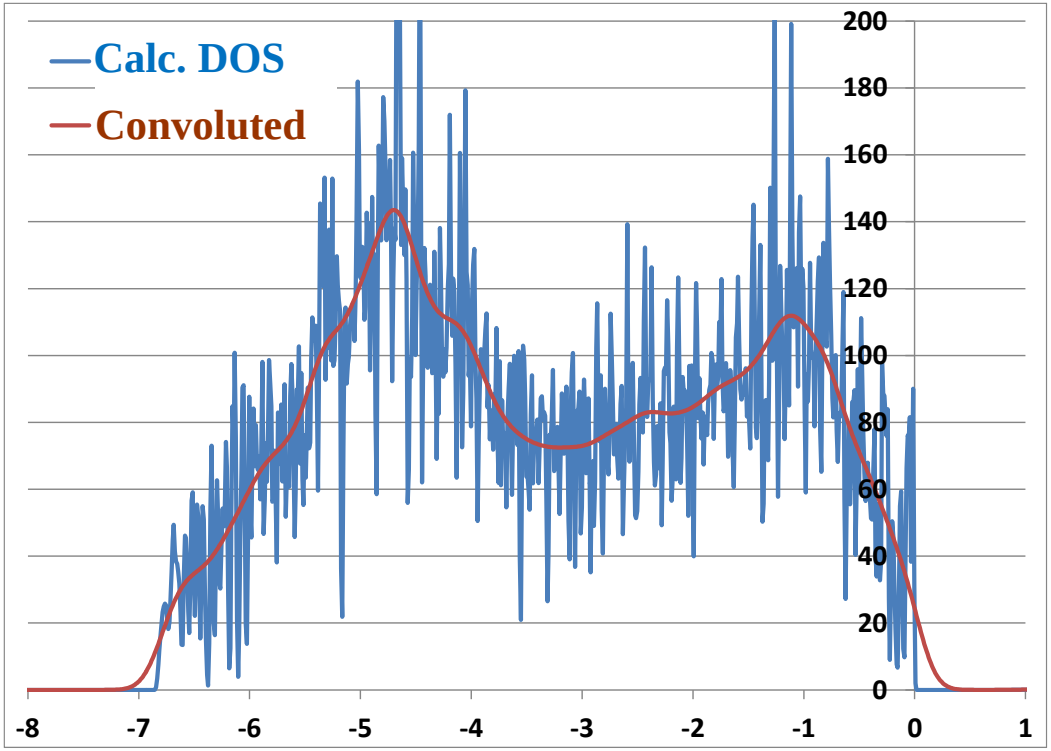
Density of state (DOS) function calculated by density functional theory

密度汎関数計算で得たa-$InGaZnO_4$の状態密度

Problem: Many noise, difficult to read

Add finite-width Gauss function to each data (それぞれのデータにGauss関数の広がり)

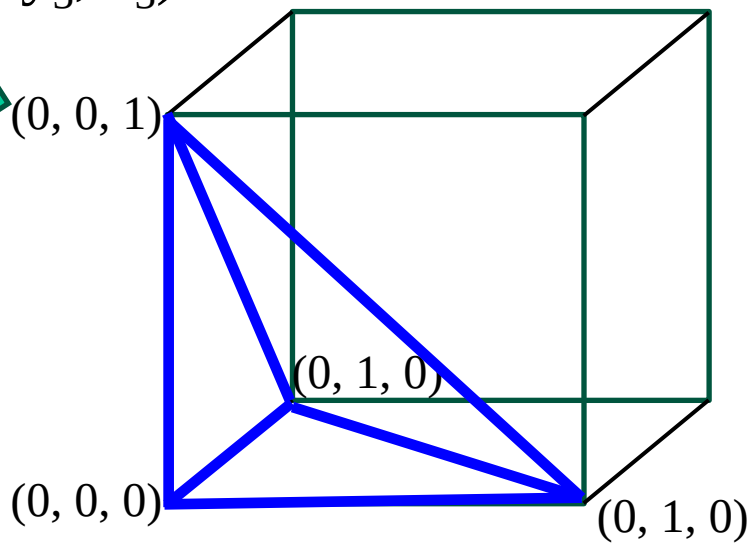
$G(E) = \exp(-[(E - E_0)/w]^2)$ (w = 0.2 eV)



Note: Estimation of band, edge energies will have the errors originating from the smearing width *w*

A: Tetrahedron method

1. Divide the first Brillouin zone to tetrahedrons
2. Choose one tetrahedron with the vertexes
   (x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)
   , normalize the vertexes to



3. Interpolate by

$$E(\boldsymbol{k}) = E_{000} + (E_{100} - E_{000})k_x + (E_{010} - E_{000})k_y + (E_{001} - E_{000})k_z$$

, where E_{ijk} is $E(\boldsymbol{k})$ at a vertex (i, j, k)

4. Integrate $E(\mathbf{k})$

No smearing of $D(E)$ => **Exact estimation of HOMO and LUMO are possible**

Q: Determination of E_V (HOMO), E_C (LUMO), E_g

1. Calculate $E(k)$ by band calculation throughout the first Brillouin zone.
2. Sort $E(k)$ from the lowest to the highest
3. Find HOMO (or E_F) level from the number of electrons (N_{tot}) and the number of orbitals included in the calculation
4. Find LUMO level as the next upper $E(k)$ from HOMO

If you have Density-Of-States data $D(E)$,

1. Integrate $D(E)$ from the lowest energy to get an integrated electron number function $N(E)$
2. Find HOMO (E_F) as the energy satisfying $N(E_F) = N_{tot}$
3. Find LUMO level as the next upper $E(k)$ from HOMO

NOTE: If $D(E)$ is smeared, it is difficult to find exact HOMO and LUMO.
Use non-smeared $D(E)$ or tetrahedron-method

For VASP, see [tkProg_Root]¥tkprog_base¥VASP¥gbandedges

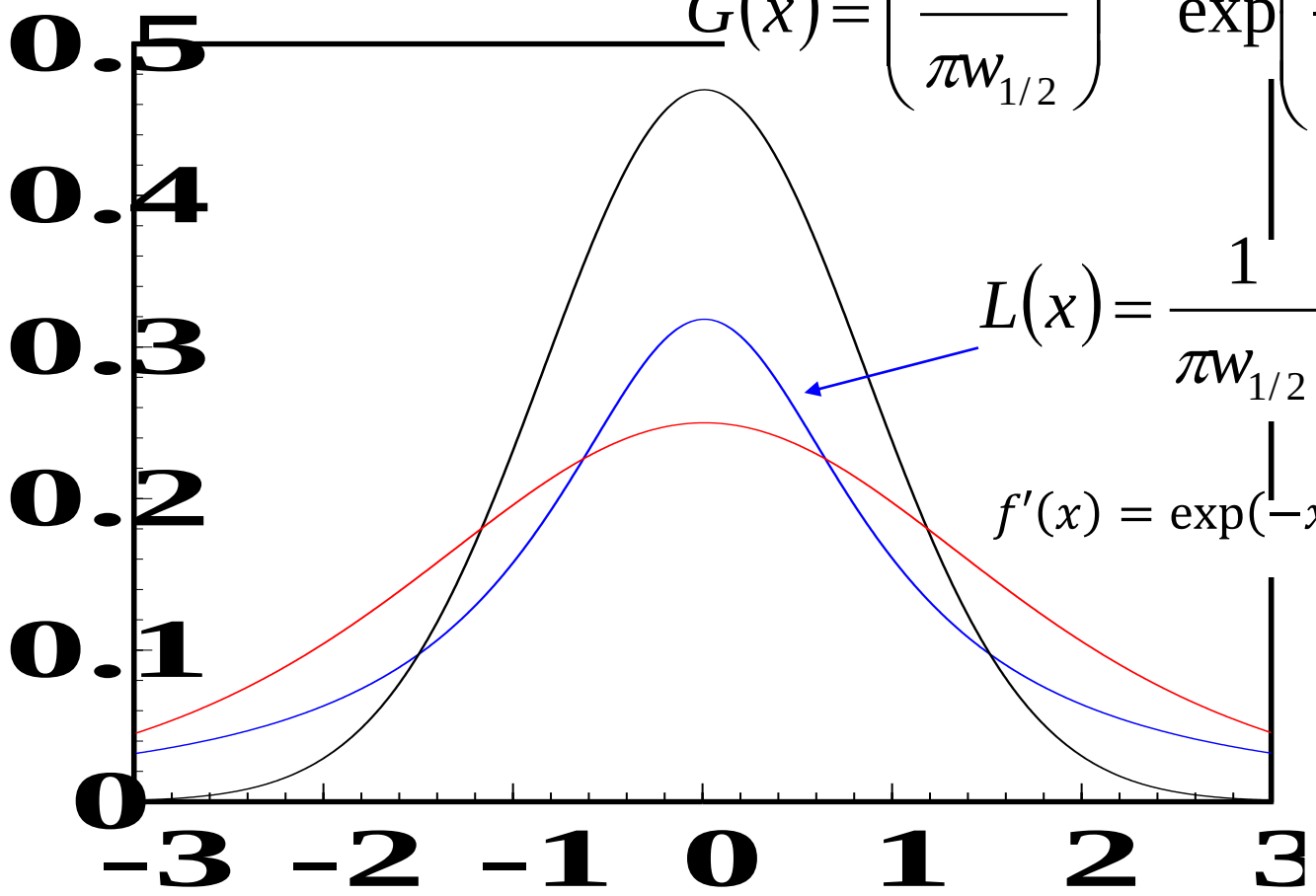
A: Smearing functions

$w_{1/2}$ = 1 eV

$$G(x) = \left(\frac{\ln 2}{\pi w_{1/2}}\right)^{1/2} \exp\left(-\ln 2\left(\frac{x}{w_{1/2}}\right)^2\right)$$

$$L(x) = \frac{1}{\pi w_{1/2}} \frac{1}{1+\left(x / w_{1/2}\right)^2}$$

$$f'(x) = \exp(-x^2)/(1+\exp(x))^2$$



A: Methfessel Paxton function

M. Methfessel and A.T. Paxton, High-precision sampling for Brillouin-zone integration in metals, Phys. Rev. B **40** (1989) 3616

Expand the delta function with Hermitian polynomials

$$\delta(x) = \sum_{n=0}^{\infty} A_n H_{2n}(x) \exp(-x^2) \qquad A_n = \frac{(-1)^n}{n! 4^n \sqrt{\pi}}$$

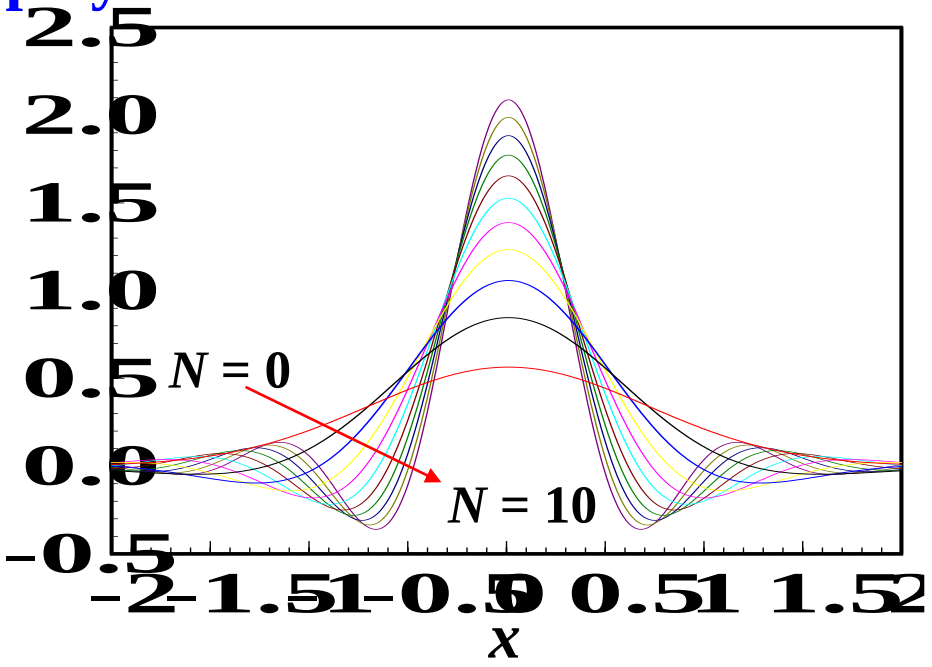
$$D_N(x) = \sum_{n=0}^{N} A_n H_{2n}(x) \exp(-x^2)$$

***$D_N(x)$* is a (2*N*+1)-order polynomial, orthogonal to a *2N* or less order polynomial**

Approximation of the Step Function

$$S_N(x) = 1 - \int_{-\infty}^{x} D_N(t)\,dt$$

$$S_0(x) = (1/2)(1 - erf(x))$$



Hermitian polynomial

$$\left(\frac{d^2}{dx^2} - 2x\frac{d}{dx} + 2n\right) H_n(x) = 0$$ **solution of a problem**

$$H_n(x) = n! \sum_{m=0}^{\mathrm{int}(n/2)} \frac{(-1)^m}{m!\,(n-2m)!} (2x)^{n-2m}$$

$$H_0(x) = 1, H_1(x) = 2x$$
$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x)$$
$$H_n'(x) = 2nH_{n-1}(x) = 2xH_n(x) - H_{n+1}(x)$$

$H_n(x)\exp(-x^2/2)$ is an orthonormal basis

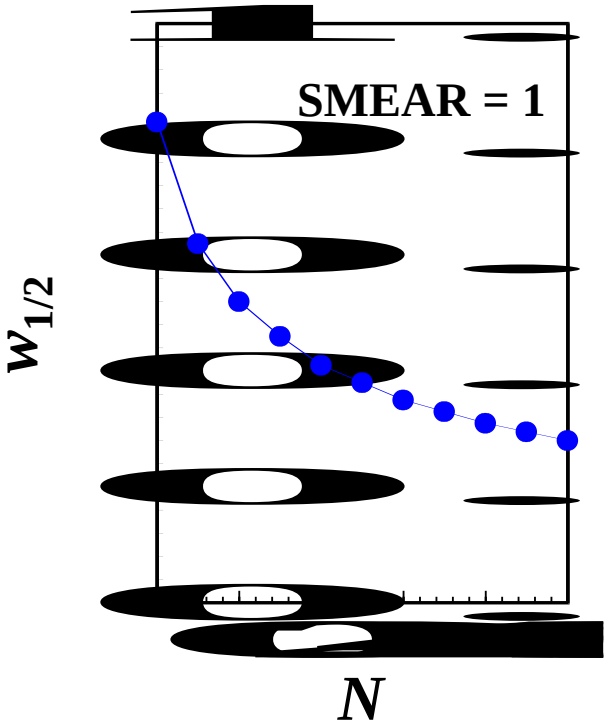
$$\int_{-\infty}^{\infty} H_n(x)H_m(x)\exp(-x^2)\,dx = \delta_{mn}2^n\sqrt{\pi}n!$$

Wavefunction of harmonic oscillator model:

$$\Psi_n(x) = (2^n\sqrt{\pi}n!)^{1/2} H_n(x)\exp(-x^2/2)$$

A: Characteristics of Methfessel Paxton Functions

1. **If the band structure and DOS can be approximated by polynomials of the *2N-th* order or less, smearing with an N-th-order MP function will not produce an integration error.**
2. **In the case of a simple band structure or DOS, integration error will be zero even if the SMEAR width is quite large.**
3. **When the band structure is complex (e.g., d-system), the optimal SMEAR width is comparable to Gaussian**
4. **The actual smearing width $w_{1/2}$ depends on the value of the order *N* and SMEAR width.**
5. **Negative values or values over 1.0 for occupancy**

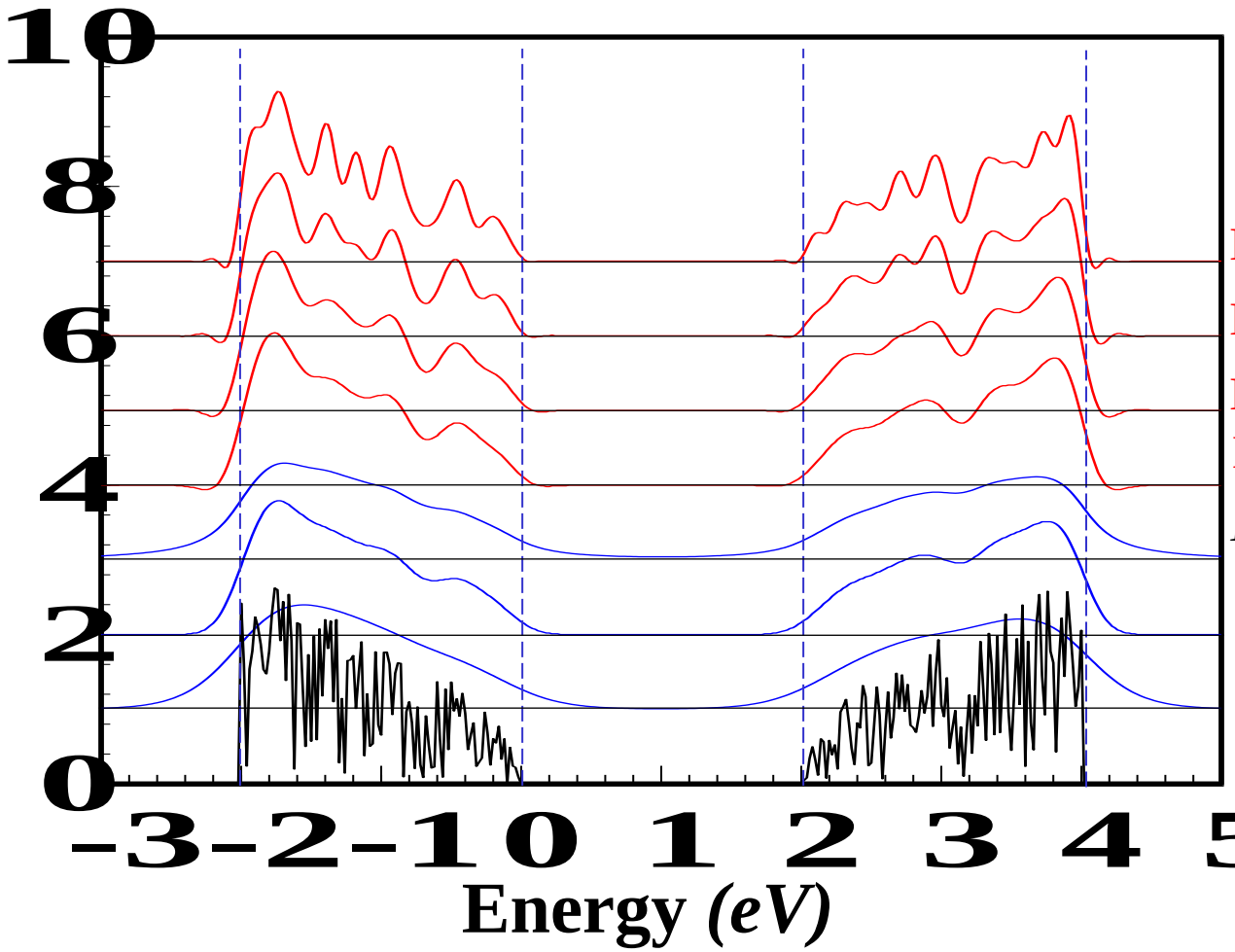


A: Smearing of density of states $D(E)*f_S(E)$

$$D(E) = D_{V0}(E_V - E)^{1/2} + D_{C0}(E - E_C)^{1/2} * (1 + \text{rand}[-0.5, 0.5])$$

Smearing: $D(E) * f_S(E) = \int D(E') f_S(E' - E) dE'$

with w = 0.2 eV



Methfessel Paxton Order 10

Methfessel Paxton Order 5

Methfessel Paxton Order 2

Methfessel Paxton Order 1

$$L(x) = \frac{1}{\pi w} \frac{1}{1 + (x/w)^2}$$

$$G(x) = (1/\pi w)^{1/2} \exp\left(-(x/w)^2\right)$$

$$f'(x) = \exp(x/w)/(1 + \exp(x/w))^2$$

$$D(E) = D_V(E) + D_C(E)$$

Q: Avoid local minimum issues

Avoid local minimum issues in non-linear optimization in particular for crystal structure determination for unknown materials

- **Examine different initial values**
- **Use algorisms that has better robustness for converging region for the initial search.**
  - **Simplex method**
  - **Simulated annealing, Dumped molecular dynamics**
    - Try to exit from the current minimum to examine another possible minimum
- **Use first-principles structure relaxation calculations to verify the stability of the analyzed structures**
- **Use structure-search programs such as**
  - **USPEX (genetic algorism) and**
  - **CARYPSO (particle swarm optimization)**

Q: Search in tree-structured data

以下の方法 に 基づいて 探索木 から 探索を 行う プログラム:
breadth-first search, best-first search, and A-star algorism

Can't give deep instructions, but for tree-structured problems:

- **Recursive programming may help to make a logic simple (but may need more memory and calculation time)**
- **Build a class that has pointers to child and parent trees**
- **For speed up, parallel search is easily applied for tree-structured problem**
  But python is not good for parallel computing due to 'Global Interpreter Lock'
  https://kosuke-space.com/python-parallel-processing

Q: Monte Carlo simulation for materials

Monte Carlo simulations:

- **Based on random number**
  How to generate random numbers in computer?
- **Application to multi-dimensional integration**
  Hit-and-miss Monte Carlo method
  Crude Monte Carlo method
- **Application to materials simulation**
  Metropolis Monte Carlo simulation

A: Monte Carlo simulation for integration

Q: 多変数の積分で、各積分範囲がお互いに別な変数を含んでしまう場合、どのように解けるのか知りたい。
Q: python scipyに付属する関数で複雑な三重積分を計算しましたが、様々なエラーによりうまくいきませんでした

A: Better to use python numpy/scipy libralies in particular for multiple integral because those are generally time-consuming calculations.

However, scipy.integrate.tplquad() is slow because it guarantees the obtained accuracy given by the argument epsrel (and .quad() for 1-D integration as well).

Therefore, the following Rieman sum may be faster that tplquad()

$$\int\int f(x,y)dxdy \sim h_x h_y \sum_i \sum_j f(x_i, y_j)$$

by Rieman sum for $x_i = x_0 + ih_x$ and $y_i = y_0 + ih_y$

If you don't need a high accuracy, Monte Calro method can be easily applied.

For infinite integrals, apply double exponential conversion algorisms

一様乱数と疑似乱数

・コンピュータで “ランダム” な事象は発生させにくい
　＝＞ アルゴリズムによって疑似乱数を発生する

・**乗積合同法**: a, b, L を正数とし、

$N_1 = a$

$N_2 = bN_1 \bmod L$　　$N \bmod L$ は N を L で除した余り

$N_3 = bN_2 \bmod L$

・・・

とすると、Nは $0 \leq N \leq L\text{-}1$ の疑似乱数になる。

・**混合合同法**: a, b, L を正数とし、

$N_1 = a$

$N_2 = bN_1 + c \bmod L$

$N_3 = bN_2 + c \bmod L$

・・・

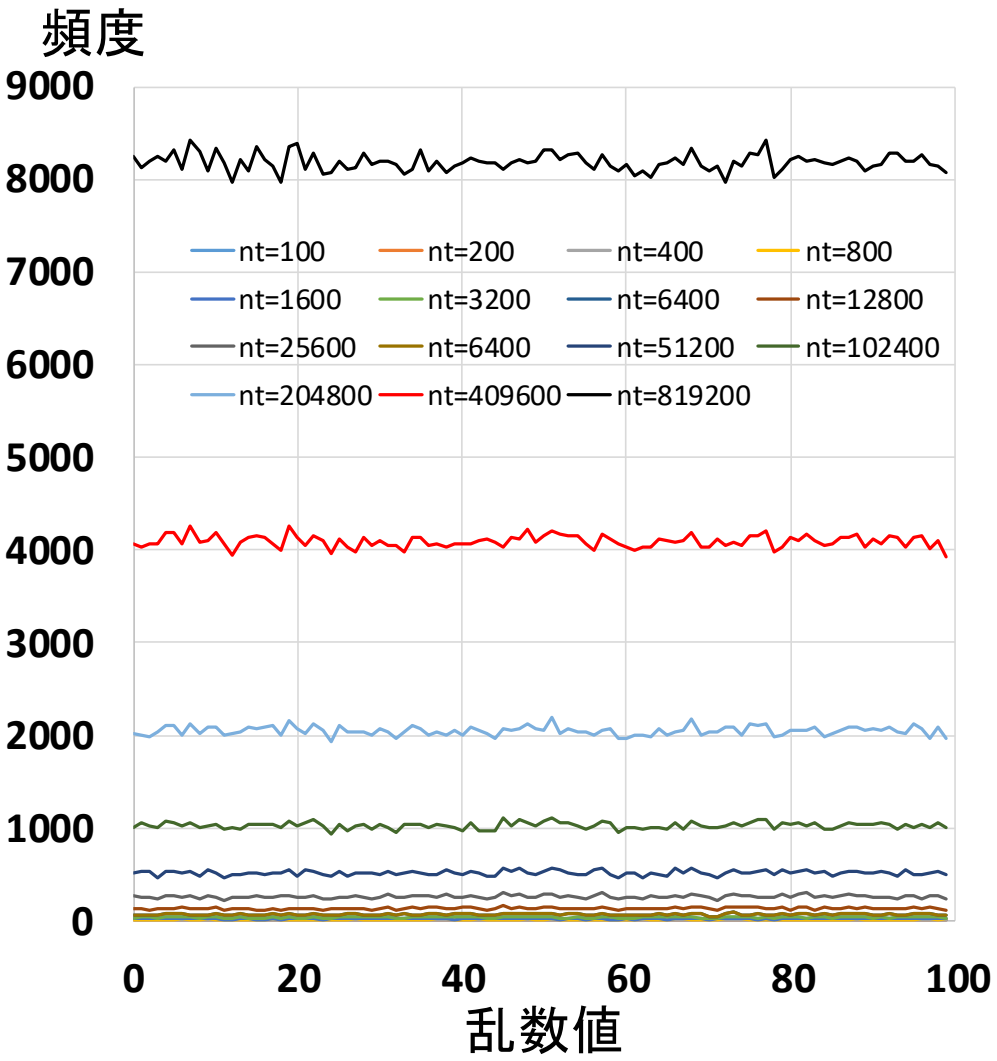
＊$N_k = N_m$ となると、乱数に周期性が発生する

疑似乱数の検定

良い疑似乱数 (一様乱数に近い) の条件

- 分布が均一
- 周期性がない
- 標準偏差が $N^{1/2}$ に比例して増大

- **疑似乱数の発生の多くには “種 (seed)” が必要**
  - seedが同じなら乱数も同じになる。
  - 毎回seedを変える必要がある
    - (時計、乱数発生器など)
- **計算ごとに乱数が変わると困ることもある**
  - (デバッグ、計算結果の比較など)
    - seedを同じにして計算



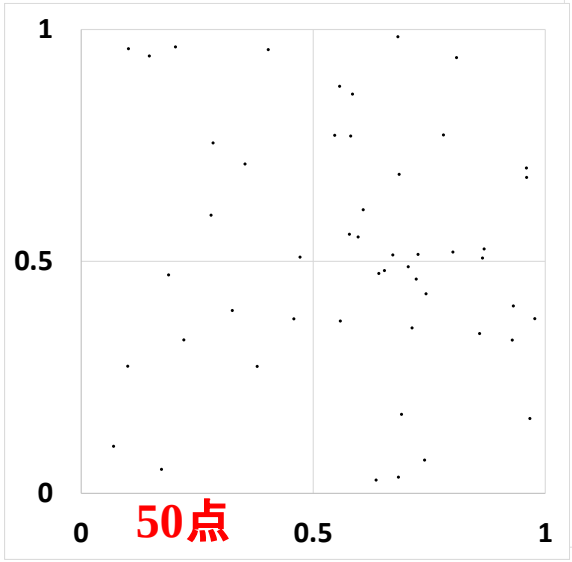
Perlの例 ＝＞

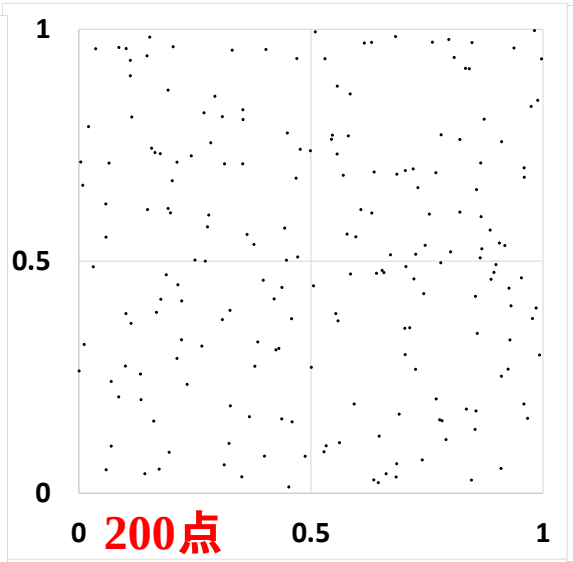
```
srand(0);
my @r;
for(my $i = 1 ; $i <= $n ; $i-
{
   $r[int(rand(100))]++;
}
```

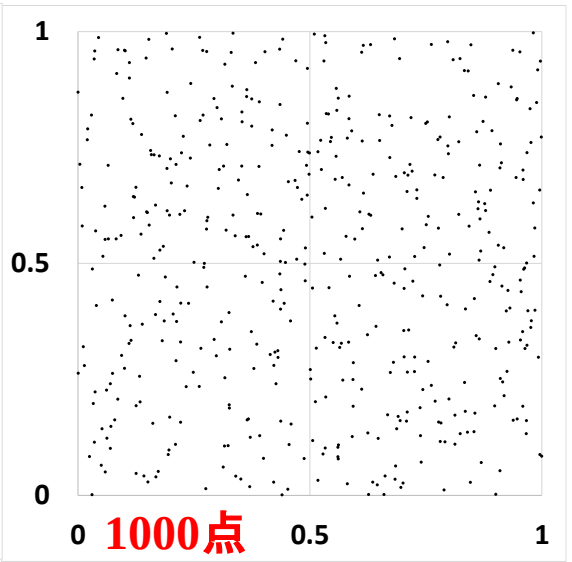
Perlのrand関数

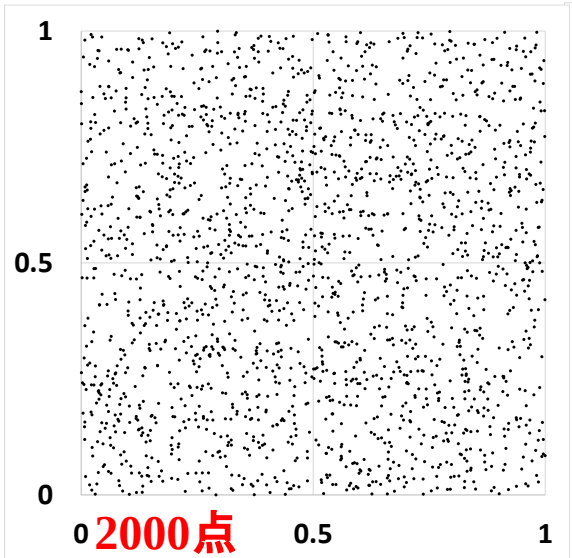
seedの指定 (srand()) は不要

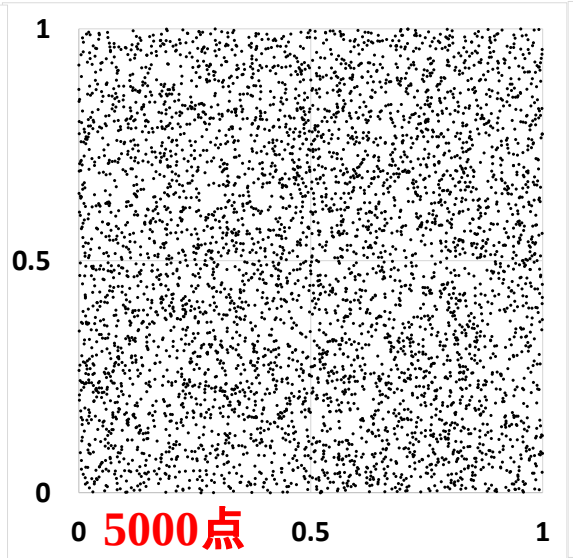
```
for(my $i = 0 ; $i < $nMax ; $i++) {
        my $x = rand(1.0);
        my $y = rand(1.0);
}
```

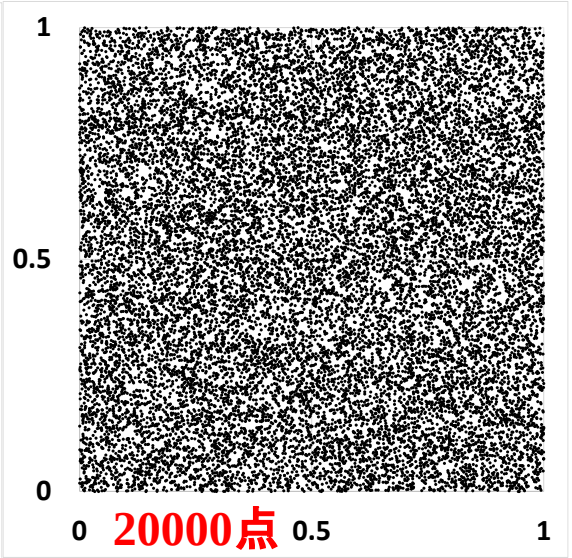






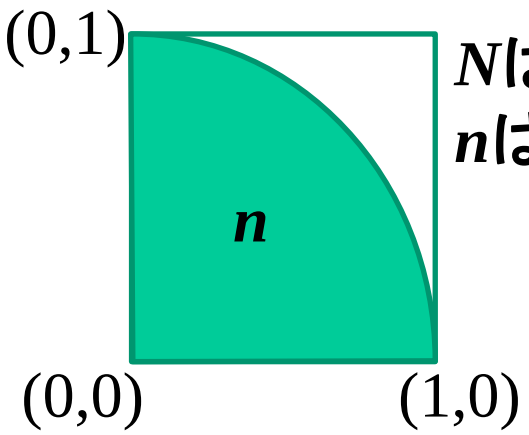






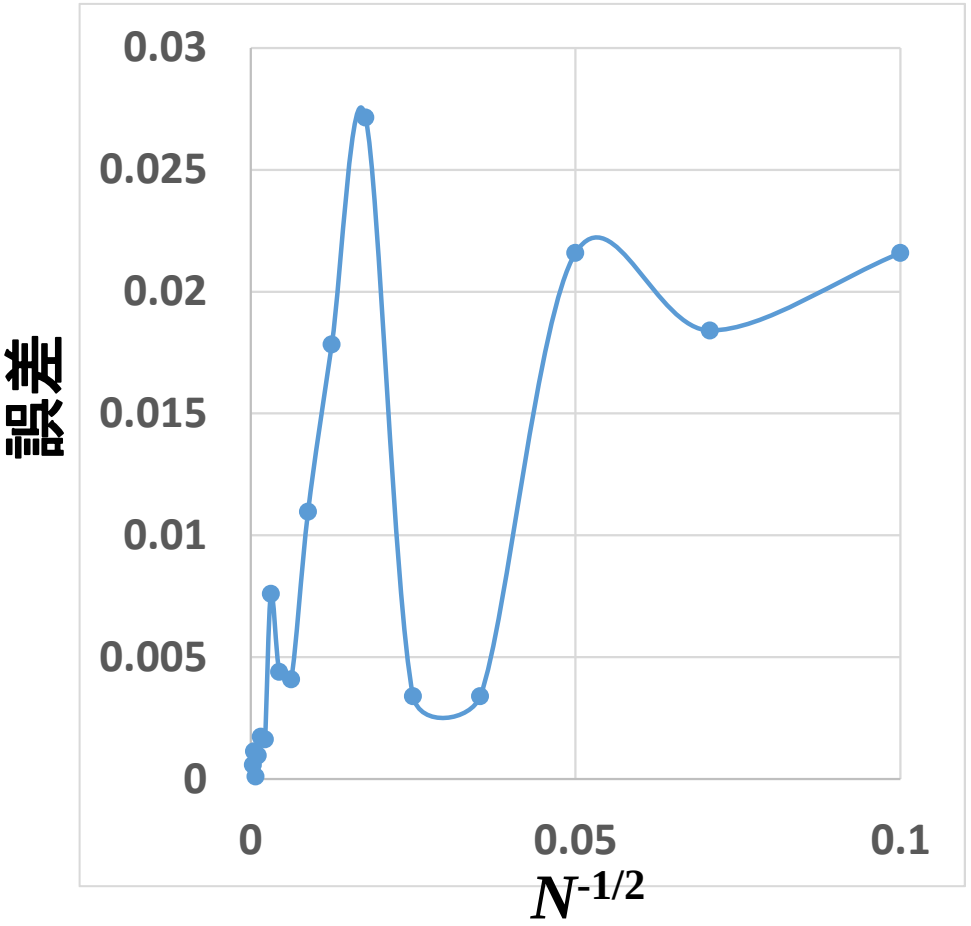
試行錯誤的 (hit-or-miss) Monte Carlo法

$0 \le r < 1$の疑似乱数 (x, y) を N回発生し、$(x^2 + y^2)^{1/2} < 1.0$ となる回数 n を求めると、n/N は 四分の一円の面積の近似値となる



Nは正方形全体に分布
nは1/4円の中に落ちた数

N	$N^{-1/2}$	4S	|error|
100	0.1	3.12	0.021593
200	0.070711	3.16	0.018407
400	0.05	3.12	0.021593
800	0.035355	3.145	0.003407
1600	0.025	3.145	0.003407
3200	0.017678	3.16875	0.027157
6400	0.0125	3.12375	0.017843
12800	0.008839	3.130625	0.010968
25600	0.00625	3.1375	0.004093
51200	0.004419	3.137188	0.004405
102400	0.003125	3.133984	0.007608
204800	0.00221	3.139961	0.001632
409600	0.001563	3.139854	0.001739
819200	0.001105	3.14063	0.000963
1638400	0.000781	3.141702	0.000109
3276800	0.000552	3.14045	0.001142
6553600	0.000391	3.141	0.000593



基礎的 (crude) Monte Carlo法

$0 \le r < 1$の疑似乱数 (x) を N回発生し、

$$S = \int_0^1 f(x)dx \sim \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

で近似できる。

$$f(x) = 4\sqrt{1-x^2}$$

N	Hit-or-mis	crude
100	2.18E-01	1.23E-01
200	1.59E-03	1.31E-02
400	1.84E-02	5.53E-02
800	3.41E-03	2.41E-02
1600	2.16E-02	1.91E-02
3200	9.66E-03	1.70E-02
6400	1.15E-02	2.69E-03
12800	9.41E-03	1.11E-03
25600	3.47E-03	1.68E-03
51200	7.69E-03	1.83E-03
102400	2.57E-03	1.95E-03
204800	5.48E-03	2.52E-03
409600	2.93E-03	9.56E-04
819200	2.50E-03	7.10E-04
1638400	4.83E-04	4.01E-04
3276800	1.62E-05	8.08E-04
6553600	1.03E-03	3.59E-04

乱数を用いた数値積分

多次元積分で用いられる

例: Discrete Variational Xα法

3次元積分点を疑似乱数で発生

integ_montecarlo3d.py

Calculate the volume of radius 1.0 sphere

Python integ_montecarlo3d.py

Output:

Hit-or-miss Monte-Carlo method

i	V	|error|
100	4.3200000000	0.13120979521360976
200	4.2000000000	0.011209795213609652
400	4.0200000000	0.16879020478639095
800	4.2000000000	0.011209795213609652
1600	4.2450000000	0.05620979521360958
3200	4.2025000000	0.013709795213609155
6400	4.1537500000	0.035040204786390916
12800	4.1868750000	0.001915204786390845
25600	4.1618750000	0.026915204786390312
51200	4.1620312500	0.026758954786390454
102400	4.1902343750	0.0014441702136096524
204800	4.1915625000	0.0027722952136093326
409600	4.1894921875	0.0007019827136094392
819200	4.1852148437	0.0035753610363906674
1638400	4.1913476562	0.002557451463609084
3276800	4.1906274414	0.0018372366198597945
6553600	4.1887829590	7.245802015276581e-06

Error ∝ 1/N

指数分布に従う乱数

http://www.sat.t.u-tokyo.ac.jp/~omi/random_variables_generation.html#Gauss

$$p(x;\lambda) = \lambda \exp\left(-\lambda x\right)$$ **(平均 1/λ, 分散 1/λ²)**

変換 *y* = exp(-*x*) を考えると、変換後の確率分布関数は

$$P(y) = P(x)\,|\,dx/dy\,|$$

となる。一様乱数 *y* から逆変換

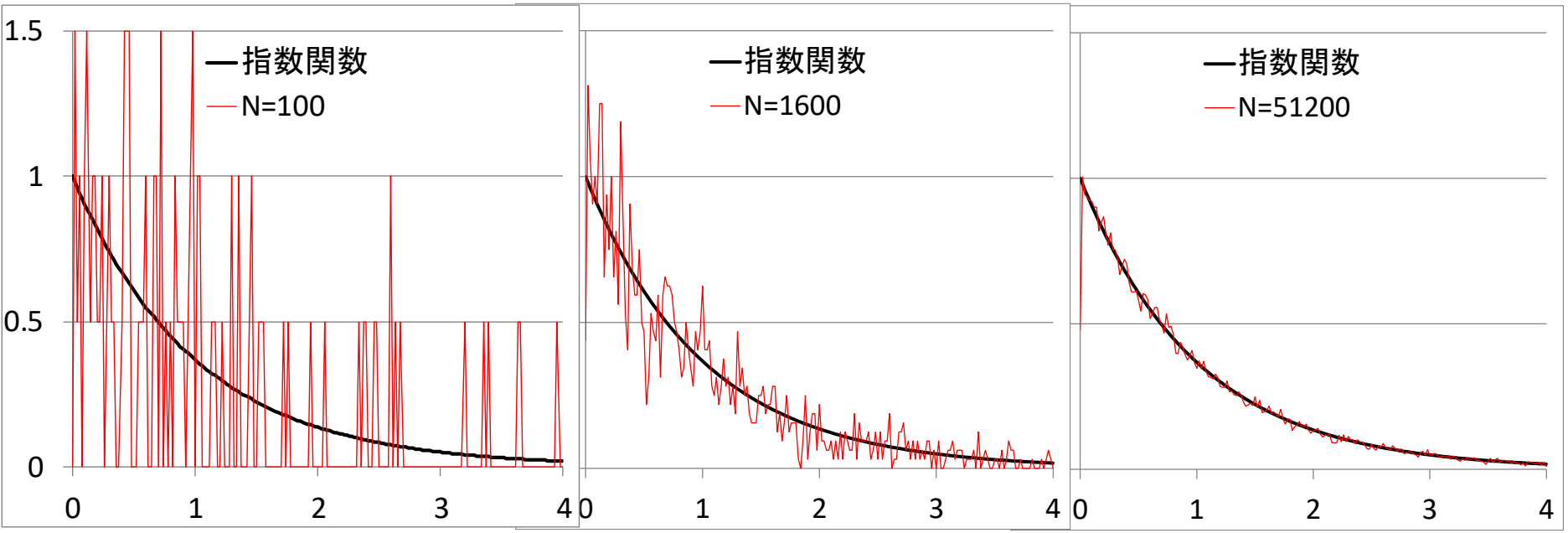
$$x = -\log(y)$$

により、λ = 1の指数分布に従う乱数が得られる。
任意の λ に対しては

$$x' = x / \lambda$$

にすればよい

指数分布に従う乱数



正規分布に従う乱数 (Box-Muller法)

http://www.sat.t.u-tokyo.ac.jp/~omi/random_variables_generation.html#Gauss

$$p(x)=\left(\frac{1}{2\pi\sigma^2}\right)^{1/2}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$ **(平均 μ, 分散 σ の正規分布)**

一様乱数 *x*, *y* を作り、極座標へ変換

$$P(x,y)=P(x)P(y)=P(r,\theta)=\left(\frac{1}{2\pi}\right)r\exp\left(-\frac{r^2}{2}\right)$$

変数を *r* から r^2 に変える $P(r^2)=P(r)\,|dx/dy|=P(r)/(2r)$

$$P(r^2,\theta)=\left(\frac{1}{4\pi}\right)\exp\left(-\frac{r^2}{2}\right)$$

一様乱数 *r*, θ から

$$x=r\cos(\theta),\ y=r\sin(\theta)$$

が正規分布に従う乱数となるので、

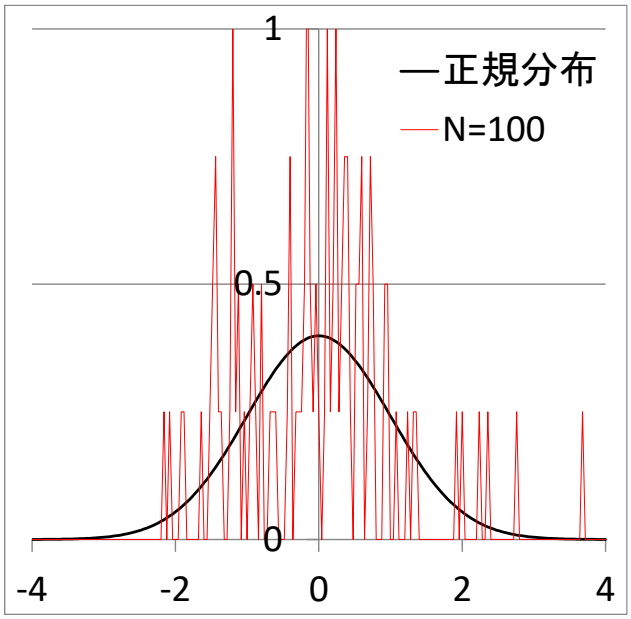
$$z=\left(-2.0*\log(x)\right)^{1/2}*\sin(2\pi y)$$

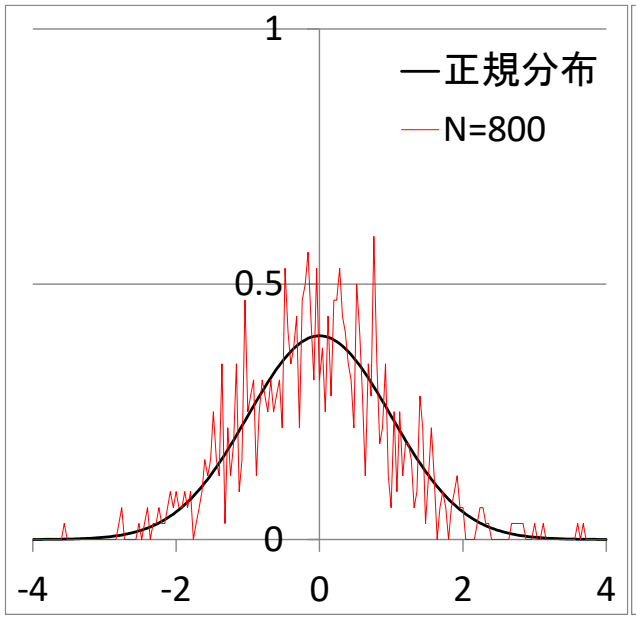
で計算できる。平均 μ, 分散 σ にするには

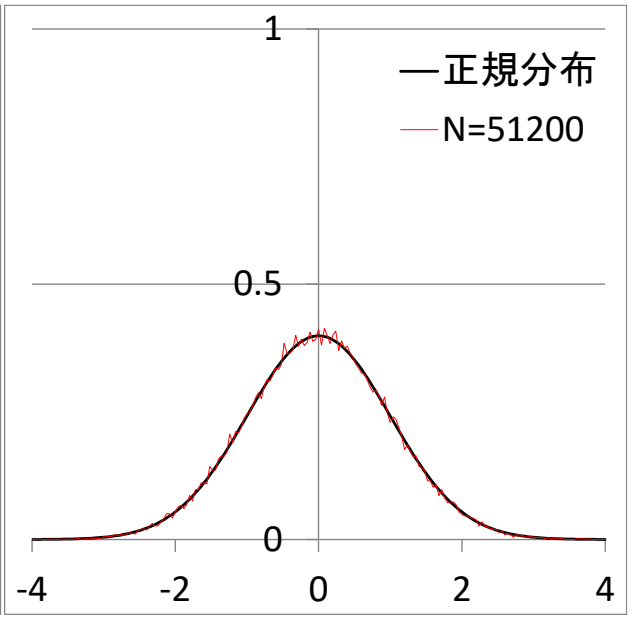
$$z' = \mu + \sigma z$$

にすればよい

正規分布に従う乱数


1
0.5
0
-4
-2
0
2
4
—正規分布
N=100


1
0.5
0
-4
-2
0
2
4
—正規分布
N=800


1
0.5
0
-4
-2
0
2
4
—正規分布
N=51200

Q: Monte Carlo simulation for materials

Question of Monte Carlo simulations for statistical physics

How to collect **an ensemble that follows**

canonical statistics $P_i \propto \exp(-E_i/k_B T)$

Metropolis Monte Carlo法

ある物理状態を考え、このポテンシャルエネルギーを計算し U_1 とする。

乱数を使って別の物理状態を作り、このポテンシャルエネルギーを U_2 とする。

1. $\Delta U = U_2 - U_1 \leq 0$ であれば、無条件にその状態を採択する
2. $\Delta U > 0$ であれば、$\exp(-\Delta U/k_B T)$ の確率で採択する

2. において、乱数 $0 \leq r \leq 1$ が $r \leq \exp(-\Delta U/k_B T)$ であれば採択、
そうでなければ棄却し、状態1 をとりもどす

という手続により作られた集団は、統計力学の母集団に一致する

この母集団について物理量の平均をとれば統計平均としての
物理量が得られる。