Newton-Cotes Integration

Recall that the Newton-Cotes formulas for approximating the definite integral of a function $f(x)$ over an interval $[a, b]$ are obtained by defining a uniform partition $x_0, x_1, x_2, \ldots, x_n$ where$\Delta x = (b-a)/n$ and $x_i = a + i\Delta x$ and interpolating values of f at the nodes in neighboring intervals by polynomials. These polynomials are then integrated analytically over their respective intervals, and the results are summed to approximate the integral of f. The various Newton-Cotes formulas vary according to the degree of the interpolating polynomial. In the second-order case, we obtain *Simpson's rule* in which three nodes x_i, x_{i+1}, x_{i+2} (two neighboring sub-intervals) at a time are used to construct a second-order polynomial $p_i(x)$ which interpolates the values $f(x_i)$, $f(x_{i+1})$, $f(x_{i+2})$. This leads to the following approximation

$$\int_a^b f(x)\,dx \approx \underbrace{\sum_{i=0}^{n-2}}_{i\,\text{even}} \int_{x_i}^{x_{i+2}} p_i(x)\,dx = \underbrace{\sum_{i=0}^{n-2}}_{i\,\text{even}} \left(c_0 f(x_i) + c_1 f(x_{i+1}) + c_2 f(x_{i+2})\right)\Delta x,$$

where the constants c_0, c_1, c_2 represent a specific linear combination of the three measurements within each sub-interval. You will now derive the values of these constants (and the resulting *Simpson rule*) in two different ways.

Part I: Direct Derivation

(a) Write down the Langrange interpolating polynomial representation for the second order polynomial $p_i(x)$ that interpolates the values $f(x_i)$, $f(x_{i+1})$, $f(x_{i+2})$ at the nodes x_i, x_{i+1}, x_{i+2}. Express your answer only in terms of $f(x_i)$, $f(x_{i+1})$, $f(x_{i+2})$ and Δx and x_i by substituting $x_{i+1} = x_i + \Delta x$ and $x_{i+2} = x_i + 2\Delta x$ and then simplifying as much as possible while still maintaining three separate terms according to the Lagrange representation).

$$p_i(x) = ?$$

(b) Compute the integral of this Lagrange interpolating polynomial over its associated sub-interval $[x_i, x_{i+2}]$. Express your answer only in terms of $f(x_i)$, $f(x_{i+1})$, $f(x_{i+2})$ and Δx.

$$\int_{x_i}^{x_{i+2}} p_i(x)\,dx = ?$$

(c) Now use your expression above in order to solve for the constant coefficients c_0, c_1, c_2:

$$c_0 = ?, \qquad c_1 = ?, \qquad c_2 = ?$$

Part II: Recursive Derivation

The Simpson's rule that you have just derived is fourth-order accurate. Using only the error formula for the second order polynomial interpolants $p_i(x)$, it is very difficult to show this (instead it would seem that the error should be only third-order accurate if you follow such an approach). In order to show that the error is indeed fourth-order accurate, we need to use the Euler-Maclaurin formula derived in class. However, we derived this directly for the *trapezoid rule*, which only employs first order polynomial interpolation rather than the second order interpolants used by *Simpson's rule*. On the other hand, the structure of the this error formula allowed us to develop the Romberg integration algorithm which shows that if we correctly combine the result of the trapezoid rule for an initial interval size (which is second-order accurate) together with the results of the trapezoid rule applied once again for half the original interval size (which is also second-order accurate), the correctly combined result will fourth order accurate. Here you will show that applying these two steps of the Romberg integration algorithm leads to the Simpson rule formula (thereby showing that it has fourth order accuracy rather than just third order accuracy).

(a) If we assume n is an even number and ignore all of the odd nodes from our original partition for part I, then we obtain a new uniform (coarser) partition $x_0, x_2, x_{4,}, \ldots, x_n$ of the same interval $[a, b]$ with a sub-intervals of width $2\Delta x$. Within each of these sub-intervals, there are now only two nodes x_i, x_{i+2}. The trapezoid rule will approximate the integrals over these respective sub-interval as

$$\int_{x_i}^{x_{i+2}} f(x)\,dx \approx (a_0 f(x_i) + a_1 f(x_{i+2}))\,(2\Delta x)$$

for some constant coefficients a_0 and a_1 which can be determined by calculating the first order polynomial that interpolates the measurements $f(x_i)$ and $f(x_{i+2})$. What are these constants?

$$a_0 =?, \qquad a_1 =?$$

(b) Now if we chop each of these larger intervals in half, we get back to our original partition (which now includes the odd nodes again). If we apply the trapezoid rule separately on two of these smaller intervals (which together make up one of the previous larger intervals) we get a new approximation of the same integral in part (a) as follows:

$$\int_{x_i}^{x_{i+2}} f(x)\,dx = \int_{x_i}^{x_{i+1}} f(x)\,dx + \int_{x_{i+1}}^{x_{i+2}} f(x)\,dx \approx (b_0 f(x_i) + b_1 f(x_{i+1}) + b_2 f(x_{i+2}))\,\Delta x$$

$$b_0 =?, \qquad b_1 =?, \qquad b_2 =?$$

(c) Now combine these two different second order approximations of the integral of f over the sub-interval $[x_i, x_{i+2}]$ according to the Romberg iteration rule to obtain a new fourth order approximation of the integral over the same sub-interval in the following form

$$\int_{x_i}^{x_{i+2}} f(x)\,dx \approx (c_0 f(x_i) + c_1 f(x_{i+1}) + c_2 f(x_{i+2}))\,\Delta x$$

where the new coefficients c_0, c_1, c_2 will depend on a combination of the previous two sets of coefficients a_0, a_1 and b_0, b_1, b_2. Do not substitute the actual values of a_0, a_1 and b_0, b_1, b_2 yet, but instead express c_0, c_1, c_2 in terms of these previous coefficients.

$$c_0 = ?, \qquad c_1 = ?, \qquad c_2 = ?$$

(d) Now substitute the values of a_0, a_1 and b_0, b_1, b_2 from part (a) and (b) into your expressions in part (c) in order to obtain the actual numerical values of the coefficients c_0, c_1, c_2 (they should match the same values you got in Part I of this problem, thus demonstrating that Simpson's rule is simply a recursive application of the trapezoid rule by following the Romberg algorithm).

$$c_0 = ?, \qquad c_1 = ?, \qquad c_2 = ?$$