TMV035 Analysis and Linear Algebra B, 2005

LECTURE 2.1

In this lecture we present the logarithm and numerical quadrature. This covers AMBS Ch 29, 30.

1. The logarithm

1.1. **The construction.** We fix a number $b > 1$ and consider the initial-value problem

$$\begin{cases} u'(x) = \dfrac{1}{x}, & x \in [1, b], \\ u(1) = 0. \end{cases} \tag{1}$$

Recall that $f(x) = 1/x$ is Lipschitz continuous on $[\delta, \infty)$ for any $\delta > 0$. Therefore the Fundamental Theorem of Calculus gives a unique solution

$$u(x) = \int_1^x \frac{1}{y}\, dy = \lim_{n\to\infty} \sum_{j=1}^{i} \frac{1}{x_{j-1}^n} h_n, \quad x \in [1, b]. \tag{2}$$

where $h_n = (b-1)2^{-n}$, $x_j^n = 1 + jh_n$, if i is related to n so that $x = x_i^n$, that is, $i = (x-1)/h_n$. We have constructed a new function $u(x)$. Since b is arbitrary it is defined for all $x \geq 1$. It is called *the natural logarithm* and it is denoted

$$\log(x) = \int_1^x \frac{1}{y}\, dy, \quad x \in [1, \infty). \tag{3}$$

For $0 < x \leq 1$ we define it as the backward integral:

$$\log(x) = \int_1^x \frac{1}{y}\, dy = -\int_x^1 \frac{1}{y}\, dy, \quad x \in (0, 1]. \tag{4}$$

This function is denoted $\ln(x)$ in Swedish, French, German books. Thus, $\log(x) = \ln(x)$. It is not the same as $\lg(x) = \log_{10}(x)$, the base 10 logarithm, which we will introduce soon. In Matlab the natural logarithm is computed as

```
>> y=log(x)
```

To plot the graph on $[1, 5]$:

```
>> x=1:0.01:5;
>> y=log(x);
>> plot(x,y)
```

You can also compute and plot it with your own program `my_int.m`:

```
>> [x,y]=my_int('funk',[1,5],0,1e-2)
>> plot(x,y)
```

where `funk.m` is

```
function y=funk(x)
y=1/x;
```

We now list the properties of log.

1.2. **Domain of definition.** It is clear that $\log(x)$ is defined for $x > 0$:

$$D_{\log} = (0, \infty). \tag{5}$$

Date: November 1, 2005, Stig Larsson, Mathematical Sciences, Chalmers University of Technology.

1.3. **Derivative and initial value.** It is clear from our construction, see (1), that log is uniformly differentiable on all intervals $[a, b]$ with $a > 0$ and

$$D\log(x) = \frac{1}{x}, \quad \log(1) = 0. \tag{6}$$

Hence it is Lipschitz continuous on all such intervals $[a, b]$.

1.4. **Monotonicity.** The logarithm is strictly increasing because $D\log(x) = \frac{1}{x} > 0$ for $x > 0$. Hence $\log(x) > 0$ for $x > 1$ and $\log(x) < 0$ for $0 < x < 1$.

1.5. **Logarithm of product and quotient.**

$$\log(ab) = \log(a) + \log(b), \quad a, b > 0, \tag{7}$$

$$\log(a/b) = \log(a) - \log(b), \quad a, b > 0. \tag{8}$$

Proof. By the chain rule and (6):

$$D\log(ax) = \frac{1}{ax}\, a = \frac{1}{x},$$

which implies

$$\log(ax) = \log(x) + C.$$

We determine the constant by taking $x = 1$:

$$\log(a) = \log(1) + C = C, \quad C = \log(a).$$

Therefore

$$\log(ax) = \log(x) + \log(a).$$

With $x = b$ we get (7). The other formula (8) is obtained by using (7):

$$\log(a) = \log\left(\frac{a}{b}\cdot b\right) = \log\left(\frac{a}{b}\right) + \log(b).$$

□

1.6. **Logarithm of a power.**

$$\log(a^n) = n\log(a), \quad a > 0,\ n = 0, \pm 1, \pm 2, \ldots, \tag{9}$$

$$\log(a^r) = r\log(a), \quad a > 0,\ r \in \mathbf{Q}. \tag{10}$$

Recall that the power function x^r with rational expontent $r \in \mathbf{Q}$ was defined for $x \geq 0$ in AMBS 18.4, see Lecture 3.2 in ALA-A. We shall soon define it for $r \in \mathbf{R}$.

Proof. We first consider $r = n \in \mathbf{N}$. Repeated application of (7) with $b = a$ gives:

$$\log(a^2) = \log(a) + \log(a) = 2\log(a), \quad \log(a^n) = n\log(a).$$

Also from (8):

$$\log(a^{-1}) = \log\left(\frac{1}{a}\right) = \log(1) - \log(a) = -\log(a), \quad \log(a^{-n}) = -n\log(a).$$

Since $\log(a^0) = \log(1) = 0$ we conclude (9). For the general case we make a substitution of variables:

$$\log(a^r) = \int_1^{a^r} \frac{1}{x}\,dx = \left\{\begin{aligned} & y = x^{1/r} \\ & x = y^r \\ & \frac{dx}{dy} = ry^{r-1} \\ & dx = ry^{r-1}\,dy \end{aligned}\right\} = \int_1^a \frac{1}{y^r}\, r\, y^{r-1} dy = r\int_1^a \frac{1}{y} dy = r\log(a).$$

(Recall the derivative $\frac{dx^r}{dx} = rx^{r-1}$ from AMBS 24.10.) □

1.7. **Asymptotic behavior.**

$$\lim_{x\to\infty} \log(x) = \infty, \tag{11}$$

$$\lim_{x\to 0^+} \log(x) = -\infty. \tag{12}$$

Here $x \to 0^+$ means that $x \to 0$ with $x > 0$, i.e., x tends to zero from the right.

Proof. From (9), and since $\log(10) > 0$, we get

$$\log(10^n) = n\log(10) \to \infty \quad \text{as } n \to \infty.$$

For any $x > 1$ we can find an integer n such that $10^n \le x < 10^{n+1}$. Then

$$\log(x) \ge \log(10^n) = n\log(10).$$

We conclude (11) because $x \to \infty$ implies $n \to \infty$. For (12) we note that $x^{-1} \to \infty$ as $x \to 0^+$ so that

$$\log(x) = -\log(x^{-1}) \to -\infty \quad \text{as } x \to 0^+.$$

□

Thus $\log(x)$ tends to infinity as x tends to infinity. But the slope of the graph is $1/x$ which tends to zero. Therefore we expect that $\log(x)$ tends to infinity very slowly. In fact, it is possible to show that

$$\lim_{x\to\infty} x^{-r}\log(x) = 0 \quad \text{for any } r > 0. \tag{13}$$

(We do not prove it here.) Here $x^{-r} \to 0$ very slowly if $r > 0$ is small and $\log(x) \to \infty$, but nevertheless $x^{-r}\log(x) \to 0$. In other words: $\log(x) \to \infty$ more slowly than x^r for any $r > 0$.

1.8. **Range.** We have seen that the logarithm is strictly increasing from $-\infty$ to ∞. Since it is Lipschitz continuous it cannot skip any values. We conclude that $y = \log(x)$ takes all values y between $-\infty$ and ∞. Therefore

$$R_{\log} = \mathbf{R}. \tag{14}$$

1.9. **Graph.** With the previous information we can now draw the graph of log. It is helpful to compute tangents, for example, for $x = 0$ the derivative is 1 so the tangent is $y = x - 1$. A Matlab plot can be seen in Figure 1.

1.10. **An inequality.**

$$\log(1 + x) \le x, \quad x > -1. \tag{15}$$

The proof is left as an exercise, Problem 29.4. If we write the inequality in the equivalent form

$$\log(x) \le x - 1, \quad x > 0,$$

then Figure 1 indicates that it is true. This is not a proof, but it helps us to remember the inequality.

1.11. **Logarithmic scale.** Very large numbers and very small numbers are mapped to numbers of size near 1 by the logarithm. For example:

$$\log(10^{10}) = 10\log(10) \approx 30,$$
$$\log(10^{-10}) = -10\log(10) \approx -30.$$

This is useful when we want to graph data with the quotient of the largest to smallest number is very large. Instead of plotting $y = f(x)$ we plot $\log(y)$ as a function of $\log(x)$ if both x and y span over large ranges. This is called a log-log plot. If only one of the variables in plotted in logarithmic scale then we have a semilog plot.

For example, the power function

$$y = ax^r, \quad a, x > 0,$$

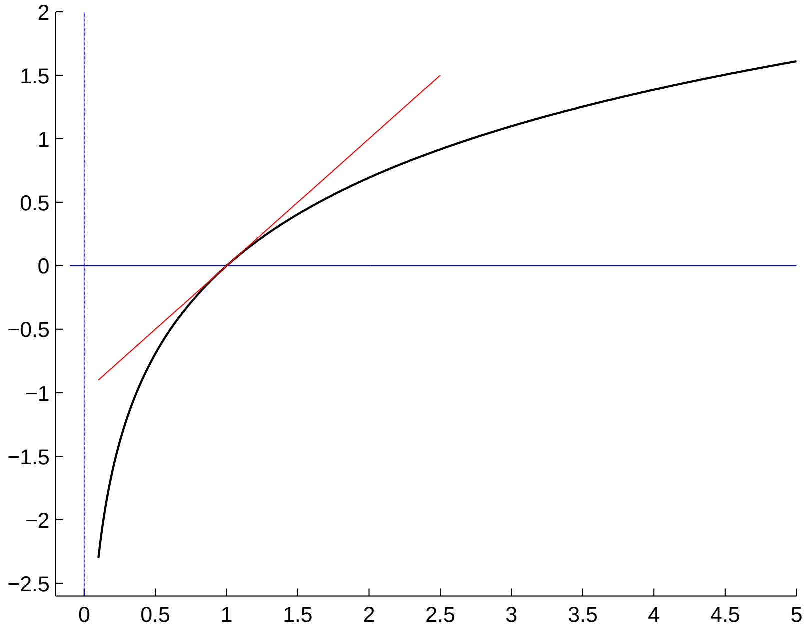

FIGURE 1. The graph $y = \log(x)$ together with the tangent $y = x - 1$.

becomes

$$\log(y) = r\log(x) + \log(a),$$

which is a straight line with slope r in the $\log(x)$-$\log(y)$ plane.

Example. We illustrate this by plotting the function $y = 10x^3$ on the interval $[10^{-3}, 10^3]$ in the original scale, see Figure 2, and in the log-log scale, see Figure 3. Note the large ranges of values spanned by both x and y: $\frac{x_{\max}}{x_{\min}} = 10^6$ and $\frac{y_{\max}}{y_{\min}} = 10^{18}$. Note that the log-log plot has slope 3 and intersects the vertical axis at $\log(10) \approx 3$.

Figure 3 was created by

```
>> x=0.001:0.1:1000;
>> y=10*x.^3;
>> plot(log(x),log(y),'*-')
>> grid on
>> axis equal
>> xlabel('log(x)')
>> ylabel('log(y)')
```

Do this now with $y = 10x^{-3}$.

MATLAB has functions for logarithmic plots: `loglog, semilogx, semilogy`. They are based on the base 10 logarithm $\lg = \log_{10}$. Figure 4 was created by

```
>> loglog(x,y)
>> grid on
```

2. Numerical quadrature

The word "quadrature" refers to computation of area. You may have heard of the classical problem of "the quadrature of the circle", which is to find a square with the same area as a given circle. We know that computation of area is more or less the same as computation of integrals.

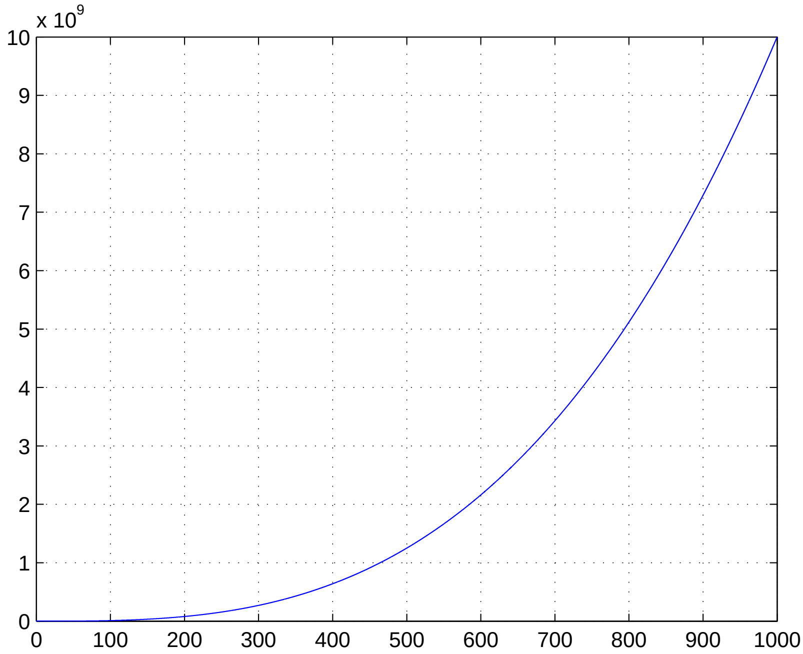

FIGURE 2. Plot of $y = 10x^3$.

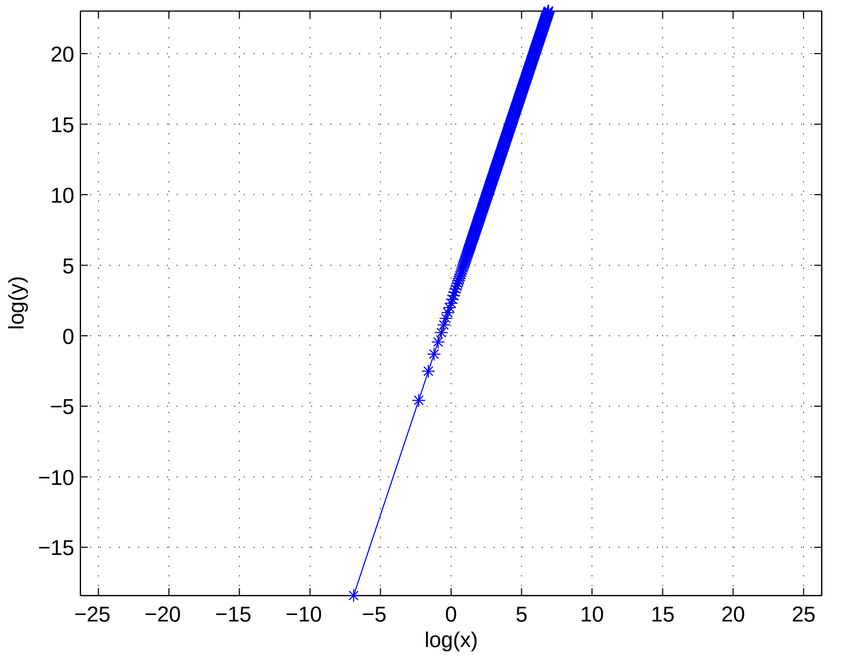

FIGURE 3. Log-log plot of $y = 10x^3$.

Therefore "numerical quadrature" means numerical computation of integrals. A "quadrature rule"

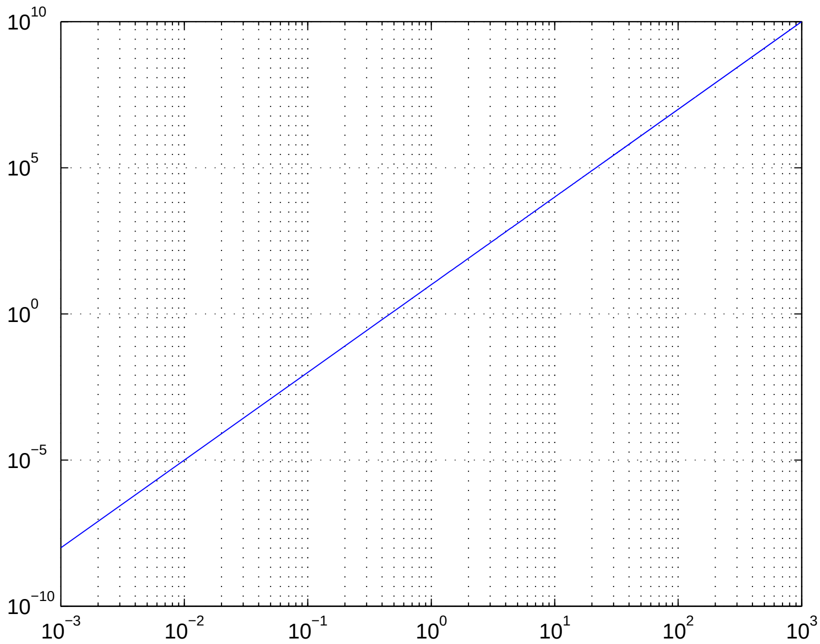

FIGURE 4. Log-log plot of $y = 10x^3$.

is an algorithm for the numerical computation of integrals. We have seen one such quadrature rule in the proof of the Fundamental Theorem of Calculus, namely the rectangle rule.

2.1. **The rectangle rule.** The Fundamental Theorem of Calculus constructs the integral as

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{j=1}^{N} f(x_{j-1}^n) h_n, \quad h_n = (b-a)2^{-n},\ x_j^n = a + jh_n,\ N = 2^n. \tag{16}$$

We shall now estimate how fast the sum on the right converges to the integral as $n \to \infty$, or equivalently, as $h_n \to 0$. But first we replace the sum by a more general quadrature rule:

- variable steplength $h_j = x_j - x_{j-1}$ and intervals $I_j = [x_{j-1}, x_j]$;
- arbitrary interpolation point $\hat{x}_j \in I_j$.

Then we get

$$\int_a^b f(x)\,dx \approx \sum_{j=1}^{N} f(\hat{x}_j) h_j. \tag{17}$$

The sum on the right is called a Riemann sum. It can also be viewed as the general form of the rectangle rule. The sum in (16) is of this form with $\hat{x}_j = x_{j-1}^n$ and $h_j = (b-a)2^{-n}$.

We now consider the quadrature error:

$$Q_h = \Big| \int_a^b f(x)\,dx - \sum_{j=1}^{N} f(\hat{x}_j) h_j \Big|. \tag{18}$$

Theorem. (Error estimate for the general rectangle rule) *Assume that f is differentiable on $[a,b]$ with Lipschitz continuous and bounded derivative. Then*

$$Q_h = \Big| \int_a^b f(x)\,dx - \sum_{j=1}^{N} f(\hat{x}_j) h_j \Big| \le \frac{1}{2} \max_{y\in[a,b]} |f'(y)| (b-a) h, \tag{19}$$

where $h = \max_{1\le j\le N} h_j$.

This means that the error converges to zero when the maximal steplength h tends to zero. The rate of convergence is

$$Q_h \le K_1 h, \quad K_1 = \frac{1}{2}\max_{y\in[a,b]} |f'(y)|(b-a). \tag{20}$$

Thus, we gain one decimal of accuracy if we decrease the steplength by a factor 0.1.

Another consequence of the theorem is that all Riemann sums

$$\sum_{j=1}^{N} f(\hat{x}_j)h_j$$

converge to $\int_a^b f(x)\,dx$ as the maximal step h goes to zero, independently of the choice of interpolation point $\hat{x}_j$ and mesh intervals I_j. This proves that the integral is unique.

Proof. By writing $\int_a^b f(x)\,dx = \sum_{j=1}^N \int_{x_{j-1}}^{x_j} f(x)\,dx = \sum_{j=1}^N \int_{I_j} f(x)\,dx$ and using the triangle inequality for sums, we get

$$Q_h = \Big|\sum_{j=1}^{N}\Big(\int_{I_j} f(x)\,dx - f(\hat{x}_j)h_j\Big)\Big| \tag{21}$$

$$\le \sum_{j=1}^{N}\Big|\int_{I_j} f(x)\,dx - f(\hat{x}_j)h_j\Big|. \tag{22}$$

This expresses the global error as the sum of the local errors. We now look at the local error:

$$\begin{aligned}\Big|\int_{I_j} f(x)\,dx - f(\hat{x}_j)\underbrace{h_j}_{=\int_{I_j}dx}\Big| &= \Big|\int_{I_j} f(x)\,dx - \int_{I_j} f(\hat{x}_j)\,dx\Big| \\ &= \Big|\int_{I_j}\big(f(x)-f(\hat{x}_j)\big)\,dx\Big| \\ &\le \int_{I_j}|f(x)-f(\hat{x}_j)|\,dx.\end{aligned}$$

Here

$$|f(x)-f(\hat{x}_j)| = \Big|\int_{\hat{x}_j}^{x} f'(y)\,dy\Big| \le \max_{y\in I_j}|f'(y)||x-\hat{x}_j|.$$

Hence

$$\Big|\int_{I_j} f(x)\,dx - f(\hat{x}_j)h_j\Big| \le \max_{y\in I_j}|f'(y)|\int_{I_j}|x-\hat{x}_j|\,dx \le \max_{y\in I_j}|f'(y)|h_j^2.$$

Here we used the fact that $|x-\hat{x}_j|\le h_j$ so that $\int_{I_j}|x-\hat{x}_j|\,dx \le h_j^2$. A more detailed calculation shows that

$$\begin{aligned}\int_{I_j}|x-\hat{x}_j|\,dx &= \int_{x_{j-1}}^{\hat{x}_j}(\hat{x}_j-x)\,dx + \int_{\hat{x}_j}^{x_j}(x-\hat{x}_j)\,dx = \frac{1}{2}\Big((\hat{x}_j-x_{j-1})^2+(x_j-\hat{x}_j)^2\Big) \\ &= \Big\{\hat{x}_j = x_{j-1}+sh_j,\ \ 0\le s\le 1\Big\} = \frac{1}{2}\Big(s^2+(1-s)^2\Big)h_j^2 \le \frac{1}{2}h_j^2.\end{aligned} \tag{23}$$

With this slightly smaller estimate we get instead

$$\Big|\int_{I_j} f(x)\,dx - f(\hat{x}_j)h_j\Big| \le \max_{y\in I_j}|f'(y)|\int_{I_j}|x-\hat{x}_j|\,dx \le \frac{1}{2}\max_{y\in I_j}|f'(y)|h_j^2. \tag{24}$$

Finally we insert this bound of the local error into (21) to get

$$Q_h \leq \frac{1}{2}\sum_{j=1}^{N} \max_{y\in I_j}|f'(y)|h_j^2 \leq \frac{1}{2}\max_{y\in[a,b]}|f'(y)|\Big(\sum_{j=1}^{N} h_j\Big)\max_{1\leq j\leq N} h_j = \frac{1}{2}\max_{y\in[a,b]}|f'(y)|(b-a)h,$$

because $\sum_{j=1}^{N} h_j = b-a$. □

2.2. **The midpoint rule.** The quantity in (23) is minimal (and $=\frac{1}{4}h_j^2$) if and only if $s=\frac{1}{2}$. This means that $\hat{x}_j = x_{j-1}+\frac{1}{2}h_j = \frac{1}{2}(x_{j-1}+x_j)$ is the midpoint of I_j. With this choice of interpolation point we get the so-called *midpoint rectangle rule* (or just the *midpoint rule*). It converges much faster than the general rectangle rule.

Theorem. (Error estimate for the midpoint rectangle rule) *Assume that f is twice differentiable on $[a,b]$ with Lipschitz continuous and bounded second derivative. Then, with $\hat{x}_j = x_{j-1}+\frac{1}{2}h_j = \frac{1}{2}(x_{j-1}+x_j)$, we have*

$$Q_h = \Big|\int_a^b f(x)\,dx - \sum_{j=1}^{N} f(\hat{x}_j)h_j\Big| \leq \frac{1}{24}(b-a)\max_{y\in[a,b]}|f''(y)|h^2, \tag{25}$$

where $h=\max_{1\leq j\leq N} h_j$.

We skip the proof. It is based on Taylor's formula.

The rate of convergence for the midpoint rule is

$$Q_h \leq K_2h^2, \quad K_2 = \frac{1}{24}(b-a)\max_{y\in[a,b]}|f''(y)|. \tag{26}$$

Thus, we gain two decimals of accuracy if we decrease the steplength by a factor 0.1.

2.3. **The trapezoidal rule.** The rectangle rule is based on integrating the piecewise constant function which is obtained by interpolating the function f in a point $\hat{x}_j\in I_j$. This gives rise to

$$\int_{I_j} f(x)\,dx \approx f(\hat{x}_j)h_j.$$

So the local integral is approximated by a small rectangle area. Another possibility is to interpolate f by the piecewise linear function:

$$f(x_{j-1})\frac{x_j-x}{h_j} + f(x_j)\frac{x-x_{j-1}}{h_j}.$$

If we integrate this we get

$$\int_{I_j} f(x)\,dx \approx \frac{1}{2}\big(f(x_{j-1})+f(x_j)\big)h_j.$$

This means that the local integral is approximated by the area of a small trapeze. For the global integral we get the *trapezoidal rule*:

$$\begin{aligned}\int_a^b f(x)\,dx &= \sum_{j=1}^{N}\int_{I_j} f(x)\,dx \approx \frac{1}{2}\sum_{j=1}^{N}\big(f(x_{j-1})+f(x_j)\big)h_j\\ &= \frac{1}{2}f(x_0)h_1 + \frac{1}{2}\sum_{j=1}^{N-1} f(x_j)\big(h_j+h_{j+1}\big) + \frac{1}{2}f(x_N)h_N.\end{aligned}$$

It can be proved that it converges with the same rate as the midpoint rule:

$$Q_h \leq Kh^2. \tag{27}$$

2.4. **Rate of convergence.** We have seen that the quadrature error satisfies

$$Q_h \leq Kh^s. \tag{28}$$

where $s = 1$ for the general rectangle rule, and where $s = 2$ for the midpoint and trapezoidal rules. Such an inequality is best illustrated in a log-log scale:

$$\log(Q_h) \leq \log(K) + s\log(h). \tag{29}$$

Here we used the fact that log is an increasing function so that the inequality is preserved. The exponent s can easily be estimated by computing the slope of the straight line.

2.5. **Adaptivity.** Recall the detailed estimate of the local error in (24):

$$\Big| \int_{I_j} f(x)\,dx - f(\hat{x}_j)h_j \Big| \leq \frac{1}{2}\max_{y\in I_j}|f'(y)|h_j^2.$$

It shows that if the derivative is large on I_j, then this term will make a large contribution to the global error. This can be compensated by choosing h_j small. And vice versa: if $\max_{y\in I_j}|f'(y)|$ is small then we can choose h_j large. So we would take small steps where f changes rapidly and large steps where f changes slowly.

An algorithm which adjusts itself to the data of the problem in this way is called an *adaptive algorithm.*

For example, if we have an error tolerance TOL, then we might choose h_j so that

$$\frac{1}{2}\max_{y\in I_j}|f'(y)|h_j \approx \frac{\text{TOL}}{b-a}.$$

Then the global error becomes

$$Q_h \leq \frac{1}{2}\sum_{j=1}^{N}\max_{y\in I_j}|f'(y)|h_j^2 = \sum_{j=1}^{N}\underbrace{\frac{1}{2}\max_{y\in I_j}|f'(y)|h_j}_{\approx\frac{\text{TOL}}{b-a}}\,h_j \approx \frac{\text{TOL}}{b-a}\Big(\sum_{j=1}^{N}h_j\Big) = \text{TOL}.$$

Further discussion of adaptivity can be found in AMBS 30.4. This is advanced, and can be skipped. `/stig`