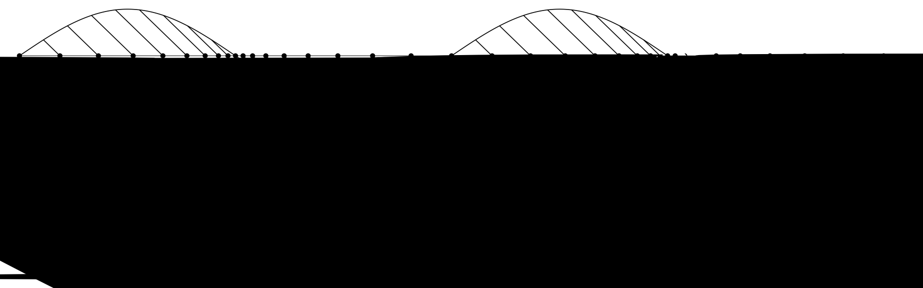
Sensible Densities

Author: Han de Bruijn
Dated: December 2004

Longitudinal Waves

In physics, there exists a method for constructing a *longitudinal wave* x_k (with variable density) from a transversal wave y_k. The geometry of this construction is rendered by the following picture:



Here Δ is a

Then the abso

Then the condition $x_{k+1} > x_k$ translates into:

$$\Delta - A \,.\, sin\left(\frac{2\pi}{\lambda}\Delta\right) > 0 \quad \Longrightarrow \quad A < \frac{\Delta}{sin(2\pi\Delta/\lambda)}$$

We can prove that a sufficient conditon is $A < \lambda/(2\pi)$, because $sin(t)/t < 1$ and hence:

$$A < \frac{\lambda}{2\pi} < \frac{\lambda/(2\pi)}{sin(t)/t} = \frac{\lambda}{2\pi}\frac{2\pi\Delta/\lambda}{sin(2\pi\Delta/\lambda)} = \frac{\Delta}{sin(2\pi\Delta/\lambda)}$$

It is concluded herefrom that: *a sufficient condition for a longitudinal wave to exist is that its amplitude shall be smaller than its wavelength λ divided by 2π.* The significance of this condition is that it is *independent* of the size of the increment Δ, which therefore can be chosen arbitrary (small).

Exact Densities

Consider a collection X of arbitrary points x_k in one-dimensional space. The members of such a *points cloud* can be thought as coordinate positions on a straight line:

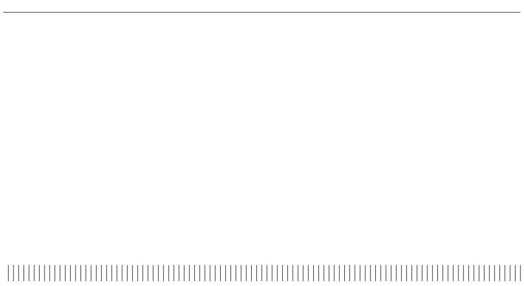
$$X = \{x_1, x_2, x_3, \ldots, x_k, \ldots, x_{n-1}, x_n, \ldots\}$$

It is assumed, in addition, that the set is ordered (sorted), in such a way that the points form a *monotone sequence*, that is:

$$j > k \quad \Longrightarrow \quad x_j > x_k$$

Now it is questioned what a *density* P is. Without doubt, by far the simplest example of a density $P(x)$ is the constant density. There will be hardly any discussion about a definition like this one:

$$P(x) = \text{constant} \quad \Longleftrightarrow \quad x_k = k\,.\,\Delta \qquad \text{where:} \quad \Delta = \text{real constant}$$



Next consider the case in which the density is increasing in a linear fashion. What does it mean: linear? Well, suppose there is one point x_k between 0 and 1, then there are two points x_k between 1 and 2, three points x_k between 2 and 3, and so on. Suppose the initial sampling distance is Δ. Then, when we arrive at $x = k.\Delta$, the number of points has been increased to $1+2+3+...+k = k(k+1)/2$. Therefore our basic equation is:

$$x_{k(k+1)/2} = L.\Delta \qquad \text{where:} \quad k, L = 1, 2, 3, ...$$

In this way, the array x is only defined for certain values of its index, namely: $1, 3, 6, ...$ In order to generalize for all values of the index, k must be solved from:

$$k(k+1)/2 = L \quad \Longrightarrow \quad k^2 + k - 2L = 0 \quad \Longrightarrow \quad k = \frac{\sqrt{8L+1} - 1}{2}$$

It is concluded that a monotone sequence with a linearly increasing density may be represented by:

$$x_L = \frac{\sqrt{8L+1} - 1}{2}\Delta \quad \text{or} \quad x_k = \left(\sqrt{1/4 + 2k} - 1/2\right)\Delta$$

The first few values (for $\Delta = 1$):

$x_0 = 0\,,\; x_1 = 1\,,\; x_2 = 1.56155\,,\; x_3 = 2\,,\; x_4 = 2.37228\,,\; x_5 = 2.70156\,,\; x_6 = 3\,,\; ...$

Other *exact densities* can be constructed by demanding that the integral over the density P, from x_k to x_{k+1}, is equal to unity:

$$\int_{x_k}^{x_{k+1}} P(t)\,dt = 1$$

Simple theorem:

$$\int_{x_0}^{x_k} P(t)\,dt = k$$

Proof by complete induction:

$$\begin{aligned}
&\int_{x_0}^{x_1} P(t)\,dt = 1 \\
&\int_{x_0}^{x_k} P(t)\,dt = k \quad \Longrightarrow \\
&\int_{x_0}^{x_{k+1}} P(t)\,dt = \int_{x_0}^{x_k} P(t)\,dt + \int_{x_k}^{x_{k+1}} P(t)\,dt = k+1
\end{aligned}$$

Specify for the *constant density*, which is given by $P(x) = 1/\Delta$, where Δ is a uniform sampling distance:

$$\int_{x_0}^{x_k} 1/\Delta\,dt = (x_k - x_0)/\Delta = k \quad \Longrightarrow \quad x_k = x_0 + k.\Delta$$

Another example has been the *linear density*, which is given by $P(x) = Cx + D$. It is no restriction on generality if the starting value x_0 of x_k is simply put to zero here. Then the integral becomes:

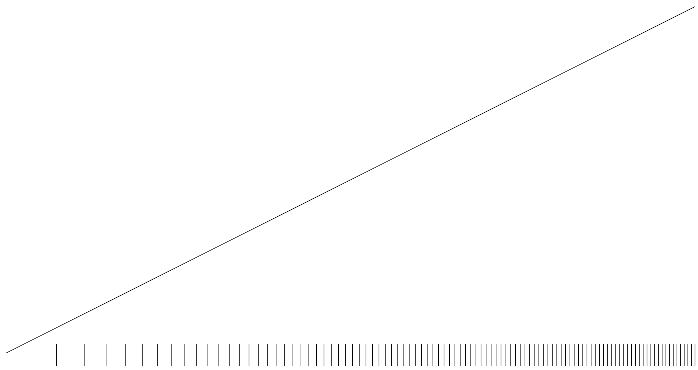
$$\int_0^{x_k} (Ct + D)\, dt = \frac{1}{2} C x_k^2 + D x_k = k \quad \Longrightarrow \quad x_k = \sqrt{(D/C)^2 + 2k/C} - D/C$$

But wait! This outcome has to be compared with the one we have obtained in an earlier stage:

$$\begin{aligned} \text{Compare} \qquad x_k &= \left(\sqrt{1/4 + 2k} - 1/2\right) \Delta \\ \text{with} \qquad x_k &= \sqrt{(D/C)^2 + 2k/C} - D/C \end{aligned}$$

A perfect match is found for $D/C = \frac{1}{2}\Delta$ and $1/C = \Delta^2$; hence $C = 1/\Delta^2$ and $D = \frac{1}{2}/\Delta$. Conclusion:

$$P(x) = \left(\frac{x}{\Delta} + \frac{1}{2}\right) / \Delta$$



A third example is the density which is given by:

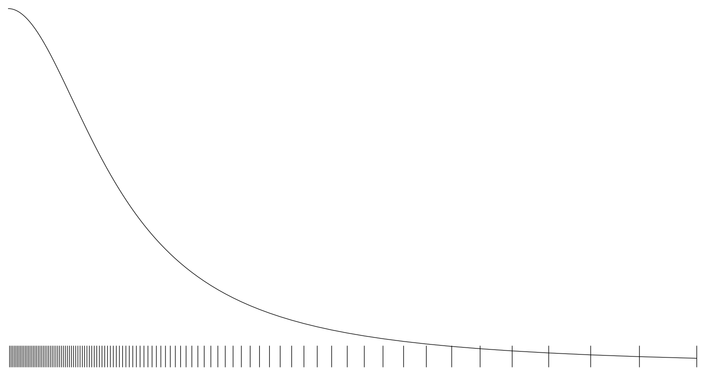
$$P(x) = \frac{1}{\Delta} \frac{d/\pi}{d^2 + x^2} = N \frac{d/\pi}{d^2 + x^2} \qquad \text{where:} \quad N = 1/\Delta$$

The area of interest is $[-\infty, +\infty]$. Resulting in:

$$N \int_{-\infty}^{x_k} \frac{d/\pi}{d^2 + t^2}\, dt = N \frac{1}{\pi} \int_{-\infty}^{x_k} \frac{d(t/d)}{1 + (t/d)^2} = N \left(\frac{1}{\pi} arctan(x_k/d) + \frac{1}{2} \right) = k$$

$$\Longrightarrow \quad arctan(x_k/d) = \frac{k}{N}\pi - \frac{1}{2}\pi \quad \Longrightarrow \quad x_k = d\ tan\left(\frac{k}{N}\pi - \frac{1}{2}\pi\right)$$

Where the possible values of k in the sequence x_k should preferrably be restricted to $0 < k < N$.



It is also possible to formulate the reverse problem: how to find the density function $P(x)$ if the sequence $\{x_k\}$ has been given. An illustrative example is the distribution of the tab stops on a guitar's neck. It can be demonstrated that it is given by:

$$x_k = L\left(\frac{1}{2}\right)^{k/12}$$

Where L is the length of the strings. The case x_0 for $k = 0$ corresponds with the full length of a string. The case $x_{12} = L/2$ for $k = 12$ corresponds with the first octave. And so on. The sequence x_k is *monotonous decreasing*. Rewrite the above formula, as follows:

$$\frac{x_k}{L} = \left(\frac{1}{2}\right)^{k/12} \quad \Longrightarrow \quad ln\left(\frac{x_k}{L}\right) = \frac{k}{12}\, ln(\frac{1}{2}) \quad \Longrightarrow \quad -12\,{}^2log\left(\frac{x_k}{L}\right) = k$$

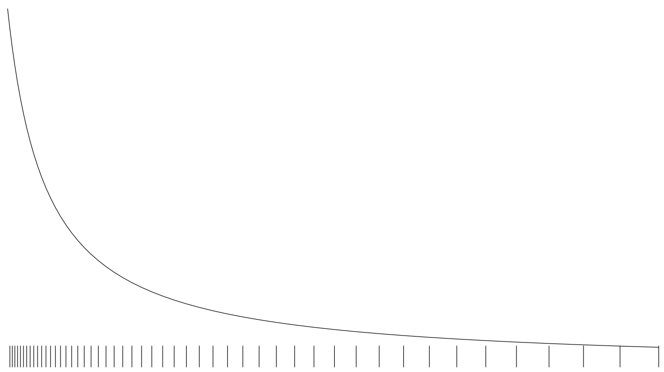
Herewith the problem is squeezed into standard form. And:

$$\int_L^{x_k} P(t)\,dt = -12\, ln\left(\frac{x_k}{L}\right) / \, ln(2) \quad \Longrightarrow \quad \int_x^L P(t)\,dt = 12\, ln\left|\frac{x}{L}\right| / \, ln(2)$$

The first minus sign stems from the fact that the integration is carried out in reverse order, from the maximal length of the string downwards to zero (due to the fact that the sequence is monotonous decreasing). By differentiating at both sides (and eventually ignoring a minus sign) we finally find:

$$P(x) = 12\,\frac{1/L}{x/L} / \, ln(2) = \frac{12}{x\; ln(2)} \sim \frac{1}{x}$$

The density function $P(x)$ is independent of the string's length L and it has a vertical asymptote for $x = 0$.



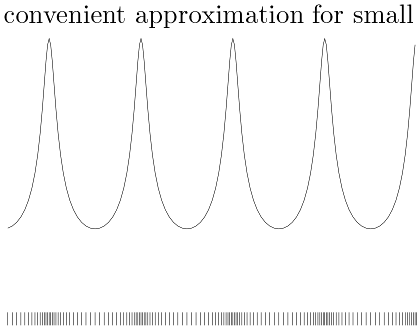
It can be demonstrated experimentally that the *density* of a longitudinal wave - contrary to what might be expected - does *not* correspond with the transversal wave from which it has been derived. Strictly spoken, it doesn't even correspond with a sinusoidal function. Nevertheless, for small amplitudes, the following approximations can be made:

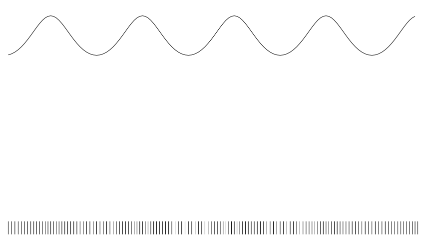
$$x_k = k \,.\, \Delta + A \,.\, sin\left(\frac{2\pi}{\lambda} k\,\Delta\right) \quad\Longrightarrow\quad k \,.\, \Delta = x_k - A \,.\, sin\left(\frac{2\pi}{\lambda} x_k\right)$$

The replacement of $(k\Delta)$ by x_k in the above equation is not at all by accident, though it can only be justified, more or less, if the amplitude A is small enough. By assuming, namely, that $k\Delta \approx x_k$, the density function can be found by differentiation:

$$\int_0^{x_k} P(t)\,dt = \frac{x_k}{\Delta} - \frac{A}{\Delta} sin\left(\frac{2\pi}{\lambda} x_k\right) \quad\Longrightarrow\quad P(x) = \frac{1}{\Delta}\left[1 - A\frac{2\pi}{\lambda} cos\left(\frac{2\pi}{\lambda} x\right)\right]$$

Giving a function $P(x)$ which is indeed maximal where the densities of the sequence x_k are most dense, as it should be. Such is the case, namely, at the odd zeroes of the accompanying transversal (sine) wave. Furthermore it is clear that densities become greater with shorter wavelengths λ. And, due to the condition $A < \lambda/(2\pi)$, the function $P(x)$ is also positive valued. It should be emphasized again, though, that the assumption $k\Delta \approx x_k$ is *not* valid in general, but only as a convenient approximation for small amplitudes.



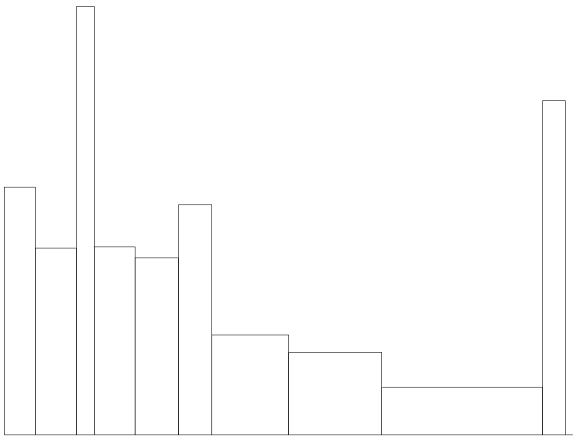


Approximate Densities

If a sequence of points $\{x_k\}$ is given, how can we find then, by a general recipe, the associated density function $P(x)$? It's not very difficult, provided that you are satisfied with a crude approximation. Consider the intervals $[x_k, x_{k+1}]$ and construct the function values in these intervals, as follows:

$$P(x) = \frac{1}{x_{k+1} - x_k} \qquad \text{for} \quad x_k \le x \le x_{k+1}$$

In such a way, namely, that the areas of the (small) rectangles, erected at the intervals $[x_k, x_{k+1}]$, are all equal to 1. With N points in $\{x_k\}$, we find $(N-1)$ "mean" function values P_k. The density, obtained by this method, is actually kind of a *histogram*:



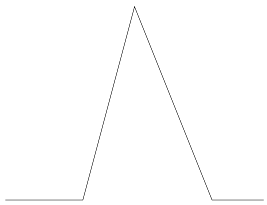
A piecewise constant distribution can also be defined as follows, which is more convenient with regard to future developments:

$$\Pi_k(x) = \begin{cases} 0 & \text{for} \quad x < x_{k-1} \\ 1/(x_{k+1} - x_k) & \text{for} \quad x_k \le x \le x_{k+1} \\ 0 & \text{for} \quad x > x_{k+1} \end{cases}$$

And:

$$P(x) = \sum_{k=1}^{N} \Pi_k(x)$$

As a next step, approximate the function values P_k by employing a piecewise *linear interpolation*. First define the following function, which has the shape of a *triangle* with base $(x_{k+1} - x_{k-1})$ and height $\Delta_k(x_k) = 1$:

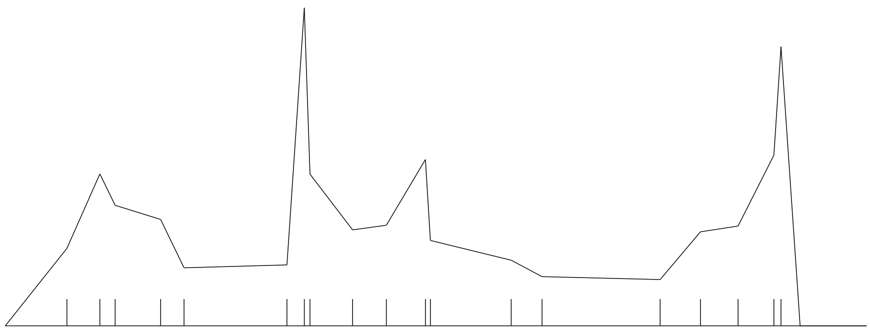


$$\Delta_k(x) = \begin{cases} 0 & \text{for} \quad x \le x_{k-1} \\ 2(x - x_{k-1})/\left[(x_k - x_{k-1})(x_{k+1} - x_{k-1})\right] & \text{for} \quad x_{k-1} \le x \le x_k \\ 2(x_{k+1} - x)/\left[(x_{k+1} - x_k)(x_{k+1} - x_{k-1})\right] & \text{for} \quad x_k \le x \le x_{k+1} \\ 0 & \text{for} \quad x \ge x_{k+1} \end{cases}$$

In such a way namely that, for each of the triangles Δ_k, the area underneath is equal to one. The base of each triangle being $(x_{k+1} - x_{k-1})$ and its height $2/(x_{k+1} - x_{k-1})$. The density function as a whole is given by:

$$P(x) = \sum_{k=1}^{N} \Delta_k(x)$$

It is remarked that $\sum_k \Delta_k(x)$ will remain zero at the end-points, that is for $k = 1$ and $k = N$. Therefore $P(x)$ will be very steep in that neighbourhood. But nevertheless, this density function is piecewise continuous everywhere:



The reverse question. Suppose that the density function $P(x)$ is given, how can we find then the associated sequence of points $\{x_k\}$? A crude approximation is to take the value $P(x_k)$ at the position x_k and construct the next point by demanding that the rectangle with height $P(x_k)$ and width $(x_{k+1} - x_k)$ has an area equal to one:

$$P(x_k)(x_{k+1} - x_k) = 1 \quad \Longrightarrow \quad x_{k+1} = x_k + \frac{1}{P(x_k)}$$

Or should the x-axis be traversed in the opposite direction? Giving:

$$P(x_k)(x_k - x_{k-1}) = 1 \quad \Longrightarrow \quad x_{k-1} = x_k - \frac{1}{P(x_k)}$$

Anyway, a better job can be done as follows. Consider a linear approximation of the density function in the neighbourhood of x_k:

$$f_{k+1} = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

Where it is noted that f_{k+1} may be almost, but not really equal to $f(x_{k+1})$. The density may be approximated with help of the trapezium rule as well:

$$\frac{1}{2}\left[f_{k+1} + f(x_k)\right]\left[x_{k+1} - x_k\right] = 1$$

Substitution of the first formula into the second gives:

$$\frac{1}{2}\left[f(x_k) + f'(x_k)(x_{k+1} - x_k) + f(x_k)\right]\left[x_{k+1} - x_k\right] = 1$$

$$\Longrightarrow \quad \left[\frac{1}{2}f'(x_k)\right](x_{k+1} - x_k)^2 + f(x_k)(x_{k+1} - x_k) - 1 = 0$$

A quadratic equation. Solve for $(x_{k+1} - x_k)$:

$$x_{k+1} - x_k = \frac{-f(x_k) \pm \sqrt{f^2(x_k) + 4f'(x_k)/2}}{f'(x_k)}$$

The numerator is of the form $(-a \pm b)$. Multiply with $(a \pm b)/(a \pm b)$, leading to values $(b^2 - a^2)/(\pm b + a)$ which is likely to be more stable numerically:

$$x_{k+1} - x_k = \frac{2f'(x_k)/f'(x_k)}{f(x_k) \pm \sqrt{f^2(x_k) + 2f'(x_k)}}$$

$$\Longrightarrow \quad x_{k+1} = x_k + \frac{1}{f(x_k)/2 + \sqrt{\left[f(x_k)/2\right]^2 + f'(x_k)/2}}$$

The plus sign (out of $\pm$) comes from the fact that, for $f'(x) = 0$, the next point has to be found by the (no longer) crude approximation: $x_{k+1} = x_k + 1/f(x_k)$. It is remarked, in addition, that the expression under the square root should not become negative. Meaning that, if $f'(x_k) < 0$:

$$|f'(x_k)| < \frac{1}{2}\left[f(x_k)\right]^2$$

It can be argued, though, that the above condition must also hold for positive values of $f'(x_k)$. The reason is that the sequence x_k can be traversed in the *opposite direction* too:

$$\left.\begin{array}{l} f_{k-1} = f(x_k) + f'(x_k)(x_{k-1} - x_k) \\ \frac{1}{2}\left[f(x_k) + f_{k-1}\right]\left[x_{k-1} - x_k\right] = -1 \end{array}\right\} \quad \Longrightarrow$$

$$\left[\frac{1}{2}f'(x_k)\right](x_{k-1}-x_k)^2+f(x_k)(x_{k-1}-x_k)+1=0$$

$$\left[\frac{1}{2}f'(x_k)\right](x_{k+1}-x_k)^2+f(x_k)(x_{k+1}-x_k)-1=0$$

Thus $(k+1)$ changes into $(k-1)$ means that $+1$ changes into -1. And that's the only difference between traversing in a positive or in a negative direction. A bit of further thinking reveals then that the negative end-result must be:

$$x_{k-1}=x_k+\frac{1}{f(x_k)/2+\sqrt{\left[f(x_k)/2\right]^2-f'(x_k)/2}}$$

And therefore, again, the abovementioned condition must hold:

$$|f'(x_k)|<\frac{1}{2}\left[f(x_k)\right]^2$$

Take for example the linear density, where $P(x)=(x/\Delta+\frac{1}{2})/\Delta$, $P'(x)=1/\Delta^2$. Then our condition translates into:

$$|P'(x_k)|<\frac{1}{2}\left[P(x_k)\right]^2 \quad\Longleftrightarrow\quad 1/\Delta^2<\frac{1}{2}\left[(x_k/\Delta+\frac{1}{2})/\Delta\right]^2 \quad\Longleftrightarrow$$

$$\left[(x_k/\Delta+\frac{1}{2})\right]^2>2 \quad\Longleftrightarrow\quad x_k/\Delta>-\frac{1}{2}+\sqrt{2} \quad\text{or}\quad x_k/\Delta<-\frac{1}{2}-\sqrt{2}$$

Only the non-negative solution must be kept. We conclude that the algoritm for constructing the set $\{x_k\}$ from the linear density function $P(x)=(x/\Delta+\frac{1}{2})/\Delta$ will only work for x_k big enough, that is: $x_k>(\sqrt{2}-1/2)\Delta\approx 0.9142135\Delta$, let's say $x_k>\Delta$.

Mean and Variance

The piecewise constant distribution has been defined as follows:

$$\Pi_k(x)=\begin{cases}0 & \text{for} \quad x<x_{k-1}\\ 1/(x_{k+1}-x_k) & \text{for} \quad x_k\le x\le x_{k+1}\\ 0 & \text{for} \quad x>x_{k+1}\end{cases}$$

A bunch of integrals will be calculated now. Start with:

$$\int_{-\infty}^{+\infty}\Pi_k(x)\,dx=\int_{x_k}^{x_{k+1}}\frac{dx}{x_{k+1}-x_k}=1$$

First order moment, center of gravity or midpoint:

$$\overline{x}=\int_{-\infty}^{+\infty}x\,\Pi_k(x)\,dx=\frac{\frac{1}{2}(x_{k+1}^2-x_k^2)}{x_{k+1}-x_k}=\frac{1}{2}(x_{k+1}+x_k)$$

Second order moment, moment of inertia or variance:

$$\overline{x^2} - \overline{x}^2 = \int_{-\infty}^{+\infty} x^2 \; \Pi_k(x) \; dx \; - \; \overline{x}^2 = \frac{\frac{1}{3}(x_{k+1}^3 - x_k^3)}{x_{k+1} - x_k} \; - \; \overline{x}^2 =$$

$$\frac{1}{3}(x_{k+1}^2 + x_{k+1}x_k + x_k^2) - \left[\frac{1}{2}(x_{k+1} + x_k)\right]^2 =$$

$$\frac{1}{3}x_{k+1}^2 + \frac{1}{3}x_{k+1}x_k + \frac{1}{3}x_k^2 - \frac{1}{4}x_{k+1}^2 - \frac{1}{2}x_{k+1}x_k + \frac{1}{4}x_k^2 =$$

$$= \frac{1}{12}(x_{k+1}^2 - 2x_{k+1}x_k + x_k^2) \quad \Longrightarrow$$

$$\overline{x^2} - \overline{x}^2 = \frac{1}{12}(x_{k+1} - x_k)^2$$

So far so good for the mean μ_k and the spread σ_k of the piecewise constant distribution:

$$\mu_k = \frac{1}{2}(x_{k+1} + x_k) \quad \text{and} \quad \sigma_k = \frac{1}{2\sqrt{3}}|x_{k+1} - x_k|$$

The piecewise linear distribution has been defined as follows:

$$\Delta_k(x) = \begin{cases} 0 & \text{for} \quad x \le x_{k-1} \\ 2(x - x_{k-1})/\left[(x_k - x_{k-1})(x_{k+1} - x_{k-1})\right] & \text{for} \quad x_{k-1} \le x \le x_k \\ 2(x_{k+1} - x)/\left[(x_{k+1} - x_k)(x_{k+1} - x_{k-1})\right] & \text{for} \quad x_k \le x \le x_{k+1} \\ 0 & \text{for} \quad x \ge x_{k+1} \end{cases}$$

A bunch of integrals will be calculated again. Start with:

$$\int_{-\infty}^{+\infty} \Delta_k(x) \; dx = \frac{1}{\frac{1}{2}(x_{k+1} - x_{k-1})} \left[\int_{x_{k-1}}^{x_k} \frac{x - x_{k-1}}{x_k - x_{k-1}} \; dx + \int_{x_k}^{x_{k+1}} \frac{x_{k+1} - x}{x_{k+1} - x_k} \; dx\right]$$

$$= \frac{1}{\frac{1}{2}(x_{k+1} - x_{k-1})} \times$$

$$\left[\frac{\frac{1}{2}(x_k^2 - x_{k-1}^2) - x_{k-1}(x_k - x_{k-1})}{x_k - x_{k-1}} + \frac{x_{k+1}(x_{k+1} - x_k) - \frac{1}{2}(x_{k+1}^2 - x_k^2)}{x_{k+1} - x_k}\right]$$

$$= \frac{\frac{1}{2}(x_k + x_{k-1}) - x_{k-1} + x_{k+1} - \frac{1}{2}(x_{k+1} + x_k)}{\frac{1}{2}(x_{k+1} - x_{k-1})} \quad \Longrightarrow$$

$$\int_{-\infty}^{+\infty} \Delta_k(x) \; dx = 1$$

First order moment, center of gravity or midpoint:

$$\overline{x} = \int_{-\infty}^{+\infty} x \; \Delta_k(x) \; dx =$$

$$\frac{1}{\frac{1}{2}(x_{k+1}-x_{k-1})}\left[\int_{x_{k-1}}^{x_k} x\ \frac{x-x_{k-1}}{x_k-x_{k-1}}\ dx + \int_{x_k}^{x_{k+1}} x\ \frac{x_{k+1}-x}{x_{k+1}-x_k}\ dx\right] =$$

$$\frac{1}{\frac{1}{2}(x_{k+1}-x_{k-1})}\times$$

$$\left[\frac{\frac{1}{3}(x_k^3-x_{k-1}^3)-x_{k-1}\frac{1}{2}(x_k^2-x_{k-1}^2)}{x_k-x_{k-1}} + \frac{x_{k+1}\frac{1}{2}(x_{k+1}^2-x_k^2)-\frac{1}{3}(x_{k+1}^3-x_k^3)}{x_{k+1}-x_k}\right] =$$

$$\frac{\frac{1}{3}(x_k^2+x_kx_{k-1}+x_{k-1}^2)-\frac{1}{2}x_{k-1}(x_k+x_{k-1})}{\frac{1}{2}(x_{k+1}-x_{k-1})} +$$

$$\frac{\frac{1}{2}x_{k+1}(x_{k+1}+x_k)-\frac{1}{3}(x_{k+1}^2+x_{k+1}x_k+x_k^2)}{\frac{1}{2}(x_{k+1}-x_{k-1})} =$$

$$\frac{\frac{1}{6}x_{k+1}^2+\frac{1}{6}x_{k+1}x_k-\frac{1}{6}x_kx_{k-1}-\frac{1}{6}x_{k-1}^2}{\frac{1}{2}(x_{k+1}-x_{k-1})} =$$

$$\frac{1}{3}\frac{(x_{k+1}+x_{k-1})\,\frac{1}{2}(x_{k+1}-x_{k-1})+x_k\,\frac{1}{2}(x_{k+1}-x_{k-1})}{\frac{1}{2}(x_{k+1}-x_{k-1})} \quad\Longrightarrow$$

$$\overline{x} = \frac{1}{3}(x_{k-1}+x_k+x_{k+1})$$

Second order moment, moment of inertia or variance. Start with:

$$\overline{x^2} = \int_{-\infty}^{+\infty} x^2\ \Delta_k(x)\ dx =$$

$$\frac{1}{\frac{1}{2}(x_{k+1}-x_{k-1})}\left[\int_{x_{k-1}}^{x_k} x^2\ \frac{x-x_{k-1}}{x_k-x_{k-1}}\ dx + \int_{x_k}^{x_{k+1}} x^2\ \frac{x_{k+1}-x}{x_{k+1}-x_k}\ dx\right] =$$

$$\frac{1}{\frac{1}{2}(x_{k+1}-x_{k-1})}\times$$

$$\left[\frac{\frac{1}{4}(x_k^4-x_{k-1}^4)-x_{k-1}\frac{1}{3}(x_k^3-x_{k-1}^3)}{x_k-x_{k-1}} + \frac{x_{k+1}\frac{1}{3}(x_{k+1}^3-x_k^3)-\frac{1}{4}(x_{k+1}^4-x_k^4)}{x_{k+1}-x_k}\right] =$$

$$\frac{\frac{1}{4}(x_k^3+x_k^2x_{k-1}+x_kx_{k-1}^2+x_{k-1}^3)-\frac{1}{3}x_{k-1}(x_k^2+x_kx_{k-1}+x_{k-1}^2)}{\frac{1}{2}(x_{k+1}-x_{k-1})} +$$

$$\frac{\frac{1}{3}x_{k+1}(x_{k+1}^2+x_{k+1}x_k+x_k^2)-\frac{1}{4}(x_{k+1}^3+x_{k+1}^2x_k+x_{k+1}x_k^2+x_k^2)}{\frac{1}{2}(x_{k+1}-x_{k-1})} =$$

$$\frac{\frac{1}{12}(x_{k+1}^3+x_{k+1}^2x_k+x_{k+1}x_k^2)-\frac{1}{4}x_k^3-\frac{1}{12}(x_{k-1}^3+x_{k-1}^2x_k+x_{k-1}x_k^2)+\frac{1}{4}x_k^3}{\frac{1}{2}(x_{k+1}-x_{k-1})} =$$

$$\frac{1}{6}(x_{k+1}^2 + x_{k+1}x_{k-1} + x_{k-1}^2) + \frac{1}{6}(x_{k+1} + x_{k-1})x_k + \frac{1}{6}x_k^2 =$$

$$\frac{1}{6}(x_{k+1}^2 + x_k^2 + x_{k-1}^2 + x_{k+1}x_k + x_{k+1}x_{k-1} + x_k x_{k-1})$$

Now substract herefrom:

$$\overline{x}^2 = \left[\frac{1}{3}(x_{k-1} + x_k + x_{k+1})\right]^2 =$$

$$\frac{1}{9}\left[x_{k+1}^2 + x_k^2 + x_{k-1}^2 + 2x_{k+1}x_k + 2x_{k+1}x_{k-1} + 2x_k x_{k-1})\right]$$

Resulting in:

$$\overline{x^2} - \overline{x}^2 = \frac{1}{18}(x_{k+1}^2 + x_k^2 + x_{k-1}^2 - x_{k+1}x_k - x_{k+1}x_{k-1} - x_k x_{k-1}) = \frac{1}{36}\times$$

$$\left[(x_{k+1}^2 - 2x_{k+1}x_k + x_k^2) + (x_{k+1}^2 - 2x_{k+1}x_{k-1} + x_{k-1}^2) + (x_k^2 - 2x_k x_{k-1} + x_{k-1}^2)\right]$$

$$\Longrightarrow \quad \overline{x^2} - \overline{x}^2 = \frac{1}{36}\left[(x_{k+1} - x_{k-1})^2 + (x_{k+1} - x_k)^2 + (x_k - x_{k-1})^2\right]$$

Hence the mean μ_k and the spread σ_k of the piecewise linear distribution:

$$\mu_k = \frac{1}{3}(x_{k-1} + x_k + x_{k+1})$$

$$\sigma_k = \frac{1}{6}\sqrt{(x_{k+1} - x_{k-1})^2 + (x_k - x_{k-1})^2 + (x_k - x_{k+1})^2}$$

Normal Densities

Consider again the sequence X of points x_k in one-dimensional space:

$$X = \{x_1, x_2, x_3, ..., x_k, ..., x_{N-1}, x_N\} \qquad \text{where} \quad (i > j) \Rightarrow (x_i > x_j)$$

As a rule, the points x_k are unevenly spaced. Nevertheless a smooth function $\overline{P}(x)$ may be associated with such a collection of points by allowing each weight m_k of a point x_k to be a continuous and differentiable function of x. We will be even more specific, though, and define:

$$m_k(x) = \frac{1}{\sigma_k\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-\mu_k)^2/\sigma_k^2}$$

And, for N points x_k:

$$\overline{P}(x) = \sum_k m_k(x) = \sum_k \frac{1}{\sigma_k\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-\mu_k)^2/\sigma_k^2}$$

Where the quantities called *mean* μ_k and *spread* σ_k are somehow related to the sequence $\{ x_k \}$. The function $\overline{P}(x)$ will be called a *Normal Density* of the sequence:

Hence the integral over any (Normal) Density happens to be exactly equal to the number of discrete points involved (in the summation):

$$\int_{-\infty}^{+\infty} \overline{P}(x)\; dx = N$$

But this property has also been encountered for the piecewise constant and the piecewise linear density. So it seems to be a universal property of density distributions. Likewise are the following:

$$\int_{-\infty}^{+\infty} x\; \overline{P}(x)\; dx = \sum_k \mu_k \quad \Longrightarrow \quad \frac{\int_{-\infty}^{+\infty} x\; \overline{P}(x)\; dx}{\int_{-\infty}^{+\infty} \overline{P}(x)\; dx} = \frac{\sum_k \mu_k}{N}$$

$$\int_{-\infty}^{+\infty} x^2\; \overline{P}(x)\; dx = \sum_k \int_{-\infty}^{+\infty} \frac{1}{\sigma_k\sqrt{2\pi}} x^2\; e^{-(x-\mu_k)^2/2\sigma_k^2}\; dx =$$

$$\sum_k \int_{-\infty}^{+\infty} \frac{1}{\sigma_k\sqrt{2\pi}} (x-\mu_k)^2\; e^{-(x-\mu_k)^2/2\sigma_k^2}\; dx \;+\; \mu_k^2 = \sum_k (\mu_k^2 + \sigma_k^2)$$

We will demonstrate now that the Normal Densities, thus defined here, are a *sensible approximation* to the Density Functions, as defined in the previous paragraphs. The following two formulas are combined in the first place:

$$\overline{P}(x) = \sum_k \frac{1}{\sigma_k\sqrt{2\pi}}\; e^{-\frac{1}{2}(x-\mu_k)^2/\sigma_k^2} \quad \text{and} \quad P\left(\frac{x_{k+1}+x_k}{2}\right)(x_{k+1}-x_k) = 1$$

$$\Longrightarrow \quad \overline{P}(x) = \sum_k \frac{1}{\sigma_k\sqrt{2\pi}}\; e^{-\frac{1}{2}(x-\mu_k)^2/\sigma_k^2} P(\mu_k)(x_{k+1}-x_k)$$

The following argument is a bit sloppy. But, for the moment being, it will be good enough. Assume that the spread σ_k is a constant σ and assume that it is large, when compared with any of the intervals: $(x_{k+1} - x_k) << \sigma$. Then μ_k can be replaced by a *continuous* variable, name it t, the increment $(x_{k+1} - x_k)$ can be replaced by a differential dt and the summation $\sum$ can be replaced by an integral $\int$. Altogether resulting in:

$$\overline{P}(x) \approx \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{1}{2}(x-t)^2/\sigma^2} P(t)\; dt = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{1}{2}t^2/\sigma^2} P(x-t)\; dt$$

We could stop here and simply state that any function $\overline{P}$, as it is a convolution of P with a normal distribution, is an approximation to the original P. However, there is a concise proof of this. Develop the function $P(x-t)$ into a power series around x:

$$P(x-t) = P(x) - t\, P'(x) + \frac{1}{2}t^2 P''(x) + ...$$

And substitute into the integral, then:

$$\overline{P}(x) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{1}{2}t^2/\sigma^2} \, P(x-t)\, dt =$$

$$P(x) \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt \quad - \quad P'(x) \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, t \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt$$

$$+ \quad \frac{1}{2} P''(x) \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, t^2 \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt \quad - \quad \ldots \quad \Longrightarrow$$

$$\overline{P}(x) \approx P(x) + \frac{1}{2}\sigma^2 . P''(x)$$

Where known properties of normal distribution functions have been employed:

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt = 1 \qquad \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, t \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt = 0$$

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \, t^2 \, e^{-\frac{1}{2}t^2/\sigma^2} \, dt = \sigma^2$$

Thus we get the result that any density function may be approximated (more or less) by a normal density function.

Fourier Transform

The function to be transformed is the normal density:

$$\overline{P}(x) = \sum_k \frac{1}{\sigma_k\sqrt{2\pi}} \, e^{-\frac{1}{2}(x-\mu_k)^2/\sigma_k^2}$$

Complex Fourier coefficients are defined as:

$$c_n = \frac{1}{T} \int_{-T/2}^{+T/2} \overline{P}(t) \, e^{-i.n.\omega.t} \, dt$$

Here T is the period of the function, which will be defined later on, i is the imaginary unit, n is an integer and $\omega = 2\pi/T$ is the angular frequency. Calculation of the Fourier coefficients will result in a Fourier series for $\overline{P(x)}$:

$$\overline{P}(x) = \sum_{n=-\infty}^{+\infty} c_n \, e^{+i.n.\omega.x}$$

Let's proceed:

$$c_n = \frac{1}{T} \int_{-T/2}^{+T/2} \left[\sum_k \frac{1}{\sigma_k\sqrt{2\pi}} \, e^{-\frac{1}{2}(t-\mu_k)^2/\sigma_k^2} \right] e^{-i.n.\omega.t} \, dt$$

$$= \frac{1}{T} \sum_k \frac{1}{\sigma_k \sqrt{2\pi}} \left[\int_{-T/2}^{+T/2} e^{-\frac{1}{2}(t-\mu_k)^2/\sigma_k^2 - i.n.\omega.t} \, dt \right]$$

The exponent of the exponential function is worked out as follows:

$$-\frac{1}{2}(t-\mu_k)^2/\sigma_k^2 - i.n.\omega.t =$$

$$-\frac{1}{2}\left(\frac{t-\mu_k}{\sigma_k}\right)^2 - i.n.\omega.\sigma_k\left(\frac{t-\mu_k}{\sigma_k}\right) - \frac{1}{2}\left(i.n.\omega.\sigma_k\right)^2$$

$$- i.n.\omega.\mu_k + \frac{1}{2}\left(i.n.\omega.\sigma_k\right)^2$$

$$= -\frac{1}{2}\left[(t-\mu_k) + i.n.\omega.\sigma_k^2\right]^2/\sigma_k^2 \; - \frac{1}{2}\left(n.\omega.\sigma_k\right)^2 - \; i.n.\omega.\mu_k \quad \Longrightarrow$$

$$e^{-\frac{1}{2}(t-\mu_k)^2/\sigma_k^2 - i.n.\omega.t} = e^{-\frac{1}{2}\left[(t-\mu_k)+i.n.\omega.\sigma_k^2\right]^2/\sigma_k^2} \; e^{-\frac{1}{2}(n.\omega.\sigma_k)^2 - i.n.\omega.\mu_k}$$

Giving:

$$c_n = \frac{1}{T} \sum_k \frac{1}{\sigma_k \sqrt{2\pi}} \; e^{-\frac{1}{2}(n.\omega.\sigma_k)^2 - i.n.\omega.\mu_k} \left[\int_{-T/2}^{+T/2} e^{-\frac{1}{2}\left[(t-\mu_k)+i.n.\omega.\sigma_k^2\right]^2/\sigma_k^2} \, dt \right]$$

In order to find a closed expression for the integral between square brackets, it will be assumed in the sequel that the interval of interest (which will be comparable to the period T of the normal density) is much greater than any of the values of the spreads σ_k:

$$\sigma_k << T$$

Herewith the integral bounds $\pm T/2$ can safely be replaced by $\pm\infty$. Giving for the integral between square brackets:

$$= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left[t-(\mu_k - i.n.\omega.\sigma_k^2)\right]^2/\sigma_k^2} \, dt = \sigma_k \sqrt{2\pi}$$

Because it makes no difference if the mean value of a normal distribution, instead of being real-valued, is actually a point in the complex plane. This outcome is cancelled against the norming factor $1/\sigma_k\sqrt{2\pi}$. So we finally find, for the Fourier coefficients of any normal density, provided that $\sigma_k << T$:

$$c_n = \frac{1}{T} \sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2 - i.n.\omega.\mu_k} = \frac{1}{T} \sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2} \; e^{-i.n.\omega.\mu_k}$$

For $n = 0$ we find:

$$c_0 = \frac{1}{T} \sum_k 1 = \frac{N}{T}$$

Therefore the accompanying Fourier series becomes:

$$\overline{P}(x) = \sum_{n=-\infty}^{+\infty} c_n \, e^{i.n.\omega.x} = \frac{1}{T} \sum_{n=-\infty}^{+\infty} \sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2} \; e^{i.n.\omega(x-\mu_k)}$$

$$= \frac{N}{T} + \frac{1}{\frac{1}{2}T} \sum_{n=1}^{\infty} \sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2} \left(e^{+i.n.\omega(x-\mu_k)} + e^{-i.n.\omega(x-\mu_k)} \right) /2$$

$$\Longrightarrow \quad \overline{P}(x) = \frac{N}{T} + \frac{1}{\frac{1}{2}T} \sum_{n=1}^{\infty} \sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2} cos\left[n.\omega(x-\mu_k)\right]$$

Where the period T still must be determined. A bit of thinking reveals that it doesn't matter much what the period looks like, as long as we know for sure that all values of the monotone sequence $\{x_k\}$ are contained in it. Now this sequence is supposed to be limited to positive values x_k. These must be in the interval $[0, \frac{1}{2}T]$. But the integration interval for the Fourier coefficients c_n originally has been defined as $[-\frac{1}{2}T, +\frac{1}{2}T]$. Meaning that values x_k in the interval $[-\frac{1}{2}T, 0]$ are needed as well. A simple manner to realize all this is to *prolongate* the sequence $\{x_k\}$ for negative $-k$ as $x_{-k} = -x_k$, declare it periodic with period T and define this period as $T = x_N - x_{-N} = 2.x_N$. Which then also becomes the period of the function $\overline{P}$. Employing this knowledge, the expression found for $\overline{P}(x)$ can be simplified even further, because:

$$\left\{ \begin{array}{l} cos\left[n.\omega(x-\mu_k)\right] = cos(n.\omega.x)cos(n.\omega.\mu_k) + sin(n.\omega.x)sin(n.\omega.\mu_k) \\ cos\left[n.\omega(x+\mu_k)\right] = cos(n.\omega.x)cos(n.\omega.\mu_k) - sin(n.\omega.x)sin(n.\omega.\mu_k) \end{array} \right.$$

Therefore the *sin* values cancel out. And we are left with the correct(ed):

$$\overline{P}(x) = \frac{N}{x_N} + \frac{2}{x_N} \sum_{n=1}^{\infty} \left[\sum_k e^{-\frac{1}{2}(n.\omega.\sigma_k)^2} cos(n.\omega.\mu_k) \right] cos(n.\omega.x)$$

Where $\omega = 2\pi/T = \pi/x_N$. N has become again the number of elements in the *original* monotone and positive sequence $\{x_k\}$, that is the one we started with. Here comes a typical picture, where the number of cosine terms is limited to 100. The discrepancy between the series and the original function at the boundaries can be explained by the fact that, at the boundaries, only *half* of the fuzzyfying function must be taken into account. This has been corrected accordingly in the accompanying program.