# Computing Elementary Functions
# Commodore 64 Style

COS:

Just add $\frac{\pi}{2}$ to the argument and fall through to SIN

SIN:

Divide argument by $2\pi$ and subtract INT from it. This sluffs off all extra cycles of $2\pi$ from the argument. Since INT rounds a negative number to the next lowest value, this in effect eliminates negative arguments too, further aiding range reduction.

Effective argument range is to be $\frac{-\pi}{2} \le \theta \le \frac{\pi}{2}$

Since division this is $-\frac{1}{4} \le \theta \le \frac{1}{4}$

If argument is greater then $\frac{3}{4}$ i.e. $\frac{3\pi}{2} \le \theta \le 2\pi$ then use identity $\sin(\theta - 2\pi) = \sin(\theta)$ , subracting one from the argument to make the range $\frac{-\pi}{2} \le \theta \le 0$

If $\frac{1}{4} \le \theta \le \frac{3}{4}$ i.e. $\frac{\pi}{2} \le \theta \le \frac{3\pi}{2}$ use identity $\sin(\pi - \theta) = \sin(\theta)$ , to subract the argument from $\frac{1}{2}$ making the range $\frac{\pi}{2} \ge \theta \ge \frac{-\pi}{2}$

Calculate the value with an odd $11^{th}$ order polynomial(6 terms) curve fitted to the function $\sin(2\pi\theta)$ (to compensate for the division) over the range $-\frac{1}{4} \le \theta \le \frac{1}{4}$ i.e. $\frac{-\pi}{2} \le \theta \le \frac{\pi}{2}$ .

LOG:

Only accept values greater then zero, a constraint imposed by logarithms.

Perform the log in base 2 to take advantage of the number format used internally by the computer. This allows the application of the formula:

$$\log_2(mantissa \times 2^{exponent}) = \log_2(mantissa) + \log_2(2^{exponent}) =$$
$$= \log_2(mantissa) + exponent$$

This is in effect a free range reduction. The log can later be converted to another base by applying the formulas:

$$\log_b x = \frac{\log_a x}{\log_a b}$$

$$\log_2 x = \frac{\log_e x}{\log_e 2} \rightarrow \log_e x = (\log_e 2) \log_2 x \quad \text{and}$$

$$\log_e x = \frac{\log_2 x}{\log_2 e} \rightarrow \log_2 x = (\log_2 e) \log_e x$$

Now all we have to do is compute the base 2 log of the mantissa which is a number guaranteed to be $1 \leq mantissa < 2$ .

We isolate the mantissa by setting the exponent of the number to 0. If the number was denormalized due to excessively small range, it will be renormalized again as this change of exponent gives the floating point routines more wiggle room.

If renormalization occurs, we siphon off any changes made to the exponent, add them to our exponent copy, and then zero out the exponent again.

Now calculate the log of the isolated mantissa. The natural series for logarithms is neither even nor odd, thus doubling the number of terms needed to calculate log to the same degree as other functions. The convergence isn't so great either.

A little manipulation is required.

We create a new variable ***t*** which is just a factorization of our original variable, but a factorization made in such a way as to enable us to split the logarithm into 2 parts, each with it's own series. Series which can be combined to cancel out alternating terms, producing a much better series.

The series are:

$$\log_e(1+t) = t - \frac{1}{2}t^2 + \frac{1}{3}t^3 - \frac{1}{4}t^4 + \frac{1}{5}t^5 - \frac{1}{7}t^7 + \cdots + \frac{1}{n}t^n \quad \text{an alternating series}$$

and

$$\log_e(1-t) = -t - \frac{1}{2}t^2 - \frac{1}{3}t^3 - \frac{1}{4}t^4 - \frac{1}{5}t^5 - \frac{1}{7}t^7 - \cdots - \frac{1}{n}t^n \quad \text{a non-alternating series}$$

(these series constrain our new variable to a range of $-1 < t < 1$
otherwise one of our log arguments could become 0 or negative)

The combination desired is:

$$\log_e(1+t) - \log_e(1-t) = 2t + \frac{2}{3}t^3 + \frac{2}{5}t^5 + \frac{2}{7}t^7 + \cdots + \frac{2}{2n-1}t^{2n-1}$$

We observe that:

$$\log_e(1+t) - \log_e(1-t) = \log_e\left(\frac{1+t}{1-t}\right) \quad \text{so} \quad \frac{1+t}{1-t}$$ must be an expression equal to our original argument

Requiring $\frac{1+t}{1-t} = x$ makes $t = \frac{x-1}{x+1}$

For the range of a normalized mantissa $1 \le x < 2$ the range of ***t*** is $0 \le t < \frac{1}{3}$ , so both our series are satisfied

However, we're not done. To minimize error we want ***t*** to be centered on zero, so both our series will expand as intended.

But only for the range $\frac{1}{2} \le x < 2$ does the ***t*** range become $-\frac{1}{3} \le t < \frac{1}{3}$ , which is symmetric alright, but that range for x is not the range of a normalized number.

We look at the powers of 2 for some guidance:

$$\frac{1}{2} \le x < 2 \rightarrow 2^{-1} \le x < 2^1$$

The powers of 2 are as symmetric as the range of ***t***. However, that range covers too many powers of 2.

Our normalized mantissa on covers one power of two:

$$2^0 \le x < 2^1$$

However, if we subtracted 0.5 from those exponents, the range would become symmetric

$$2^{-0.5} \le x < 2^{0.5}$$

Turns out, that's very easy. To add exponents together all we need to do is multiply some underlying numbers. In this case powers of 2.

$$2^x 2^y = 2^{(x+y)}$$

Choosing $2^{-0.5}$ as our factor gives:

$$\begin{gathered}\left(2^{-0.5}\right)\left(2^0\right) \le \left(2^{-0.5}\right)(x) < \left(2^{-0.5}\right)\left(2^1\right) \rightarrow \\ 2^{-0.5} < \left(2^{-0.5}\right)(x) < 2^{0.5}\end{gathered}$$ which is the desired range, and gives ***t*** a range of

$-(3-2^{1.5}) \le t < 3-2^{1.5}$ which is symmetrical about zero.

We can then incorporate the scaling factor into our formula for ***t***:

$$t = \frac{x-1}{x+1} \rightarrow t = \frac{x\left(2^{-0.5}\right)-1}{x\left(2^{-0.5}\right)+1} = \left(\frac{x\left(2^{-0.5}\right)-1}{x\left(2^{-0.5}\right)+1}\right)\left(\frac{2^{0.5}}{2^{0.5}}\right) = \frac{x-2^{0.5}}{x+2^{0.5}}$$ which replaces the

multiplication with a constant.

Now all we do is convert the series to log base 2.

$$\log_2\left(\frac{1-t}{1+t}\right)=(\log_2 e)\log_e\left(\frac{1-t}{1+t}\right)$$

The last detail is to undo the effect of multiplying the mantissa by the scaling factor

From the formula:

$$\log_2\left((2^{-.5})x\right)=\log_2(2^{-0.5})+\log_2 x=\log_2 x-0.5$$

we learn that we inadvertantly subtracted 0.5 from our answer. We're going to need to add that back.

## Putting it all together:

For $t=\frac{x-2^{0.5}}{x+2^{0.5}}$ , the base 2 logarithm of the mantissa $\boldsymbol{x}$ is:

$$\log_2 x=(\log_2 e)\log_e\left(\frac{1+t}{1-t}\right)+.5$$

Adding in our exponent:

$$\log_2\left(x\times 2^{exponent}\right)=(\log_2 e)\log_e\left(\frac{1+t}{1-t}\right)+.5+exponent$$

And finally converting to a base e log:

$$\log_e\left(x\times 2^{exponent}\right)=\left[(\log_2 e)\log_e\left(\frac{1+t}{1-t}\right)+.5+exponent\right](\log_e 2)$$

The C64 uses a $7^{th}$ order polynomial(4 terms) to evaluate.

Addendum:
If we curve fit the series to:

$(\log_2 e)\log_e\left(\frac{1+t}{1-t}\right)$ instead of $\log_e\left(\frac{1+t}{1-t}\right)$

the conversion factor $\log_2 e$ will get incorporated into the coefficients.

EXP:

For EXP we consider the form of the solution, rather then the argument when considering argument reduction.

$$e^{arg}=answer\times 2^n$$

Taking the log of both sides gives:

$arg = \log(answer \times 2^n) = \log(answer) + \log(2^n) = \log(answer) + n\log(2)$
which means our argument can be factored into:
$arg = x + n\log(2)$ which looks a lot like $arg = \theta + n2\pi$ as used in SIN

Simply subtract off the extra ***log(2)***'s.

However we can't just cast them away like we did with SIN, because they're a part of our answer.

$$e^{x+n\log(2)} = e^x e^{n\log(2)} = e^x e^{\log(2^n)} = e^x 2^n$$

***n*** is actually the mantissa of the answer. A good test for overflow is to check if this number is in the range of an exponent.

So:
$n = \mathrm{INT}\left(\frac{arg}{\log(2)}\right)$ and $x = \frac{arg}{\log(2)} - n$ and just like in SIN, INT rounds to the next lowest integer when the argument is negative, so x will always be positive.

x is now in the range $0 \le x < 1$ i.e. $0 \le x < \log(2)$

Effective argument range is to be $\frac{-\log(2)}{2} \le x \le \frac{\log(2)}{2}$ . Since division this would be $-\frac{1}{2} \le x \le \frac{1}{2}$ .

If $\frac{1}{2} \le x < 1$ then subtract 1 from x and add 1 to n to make the range $-\frac{1}{2} \le x < 0$

Curve fit the series to $e^{x\log(2)}$ to compensate for dividing by log(2)

$e^{arg} = e^{x+n\log(2)} = series \times 2^n$
The series is absolutely atrocious, neither even nor odd. The C64 computes this series with an $8^{th}$ order polynomial (using 9 terms).

Addendum:
Major improvement since C64:

Instead of factoring the argument to obtain a better series, we factor the answer.

$$e^x=(e^x-1)+1=\frac{2x}{2x}(e^x-1)+1=\frac{2x}{1}\frac{(e^x-1)}{2x}+1=\frac{2x}{\frac{2x}{e^x-1}}+1$$

$$=\frac{2x}{x\left(\frac{2}{e^x-1}\right)}+1=\frac{2x}{x\left(\frac{e^x+1+1-e^x}{e^x-1}\right)}+1=\frac{2x}{x\left(\frac{e^x+1-(e^x-1)}{e^x-1}\right)}+1$$

$$=\frac{2x}{x\left(\frac{e^x+1}{e^x-1}\right)-x}+1$$

The series

$$x\left(\frac{e^x+1}{e^x-1}\right)=2+\frac{x^2}{6}-\frac{x^4}{360}+\frac{x^6}{15120}-\frac{x^8}{604800}+\frac{x^{10}}{23950080}+\cdots$$

is a much nicer even series using half the terms for any given order. A division is added in post processing which may not save anymore time then staying with the orinal series. But the C64 only uses a 32 bit mantissa. For larger mantissa's the savings really adds up.

Argument reduction is the same.
But it would be better to multiply the “normalized” ***x*** by log(2) to get it back to the true range $-\log\frac{(2)}{2}\leq x<\log\frac{(2)}{2}$ , to simplify post processing.

So we curve fit the series to the function.

$$x\left(\frac{e^x+1}{e^x-1}\right)$$

and

$$e^{arg}=e^{x-n\log(2)}=\left(\frac{2x}{series-x}+1\right)\times 2^n$$

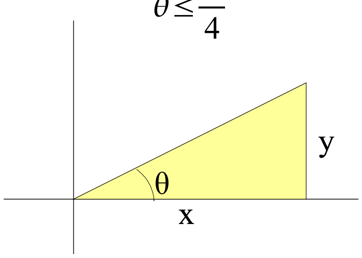
which one can evaluate with a $6^{th}$ order even polynomial (4 terms)

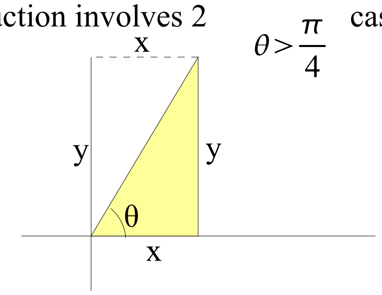
TAN:

Just compute SIN and divide by COS

ATN:

Make positive and save sign for later.
Then $\theta\leq\frac{\pi}{4}$ argument reduction involves 2 $\theta>\frac{\pi}{4}$ cases:



When $\theta \le \frac{\pi}{4}$ then $\frac{y}{x} \le 1$ and we go on to calculate ATN $\left(\frac{y}{x}\right)$

When $\theta > \frac{\pi}{4}$ then $\frac{y}{x} > 1$ and we go on to calculate ATN $\left(\frac{x}{y}\right)$ and the answer is then

$$ATN\left(\frac{y}{x}\right) = \frac{\pi}{2} - ATN\left(\frac{x}{y}\right) = \frac{\pi}{2} + ATN\left(-\frac{x}{y}\right)$$

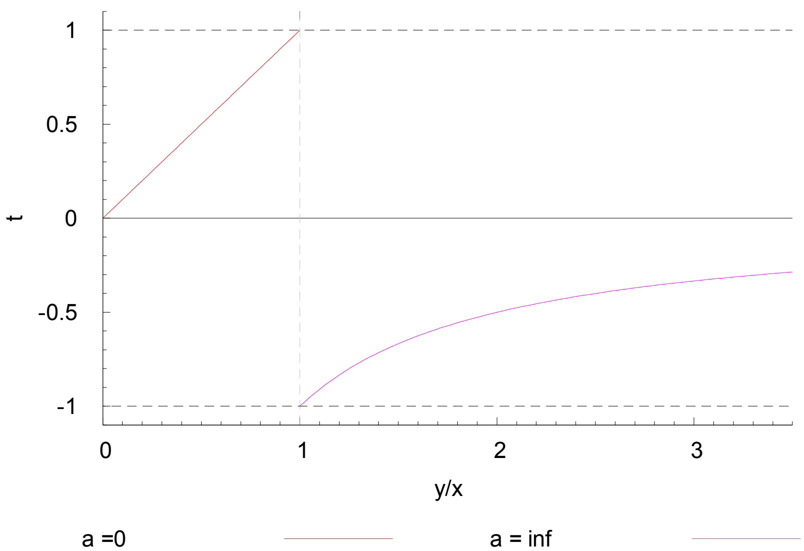
These steps make sure the argument to the series is always $0 \le arg \le 1$ . The C64 uses an 23[th] order odd polynomial (12 terms) to calculate.

Final Step:

Slap on the sign of the orginal argument to get the answer in the range $\frac{-\pi}{2} < ATN < \frac{\pi}{2}$

Addendum:
Major improvement since C64:

The C64 broke the ATN argument into two ranges:




$$ATN\left(\frac{y}{x}\right) = ATN(t), \quad t = \begin{Bmatrix} \frac{y}{x}, & \frac{y}{x} \le 1 \\ -\frac{x}{y}, & \frac{y}{x} > 1 \end{Bmatrix}$$

Taking our que from the formula:

$ATN\left(\frac{y}{x} > 1\right) = \frac{\pi}{2} + ATN\left(-\frac{x}{y}\right)$ we see if we might be able to split up the argument into more intervals.

We generalize the formula to:

$ATN\left(\frac{y}{x}\right)=ATN(a)+ATN(t)$ hoping to get a range reduced formula for ***t*** for various values of ***a***

Then we can compute the value of ATN(t) and just add a constant for the final answer.

Making use of the formula: $\tan(\alpha+\beta)=\frac{\tan(\alpha)+\tan(\beta)}{1-\tan(\alpha)\tan(\beta)}$ we take the TAN of both sides:

$$\tan\left(ATN\left(\frac{y}{x}\right)\right)=\frac{\tan(ATN(a))+\tan(ATN(t))}{1-\tan(ATN(a))\tan(ATN(t))}$$

$$\frac{y}{x}=\frac{a+t}{1-at}$$

giving us our reduction formula $t=\frac{\frac{y}{x}-a}{1+a\frac{y}{x}}$

You'll notice from the chart above, that our 2 range scheme produced a symmetric range of $-1\le t\le 1$ for ***t***. We desire a symmetric range for our new scheme as well.

Let's decide on a 3 range scheme first:

$$ATN(s)=ATN(a(s))+ATN(t(s)),\ \ t(s)=\frac{s-a}{1+a\,s},\ \ a(s)=\begin{Bmatrix}0 & 0\le s\le g1\\ k & g1\le s\le g2\\ inf & g2\le s\end{Bmatrix}$$

To get a symmetric range $-r\le t\le r$ we require:

$$\begin{aligned}t(g1^-)&=r\\ t(g1^+)&=-r\\ t(g2^-)&=r\\ t(g2^+)&=-r\end{aligned}$$

This gives us a system of 4 nonlinear equations:

$$\begin{matrix}\frac{g1-0}{1+0\,g1}=r\\ \frac{g1-k}{1+k\,g1}=-r\\ \frac{g2-k}{1+k\,g2}=r\\ \frac{g2-inf}{1+inf\ g2}=-r\end{matrix} \quad = \quad \begin{matrix}g1=r\\ \frac{g1-k}{1+k\,g1}=-r\\ \frac{g2-k}{1+k\,g2}=r\\ -\frac{1}{g2}=-r\end{matrix}$$

in 4 unknowns (***r, g1, g2, k***)

Solving gives us:

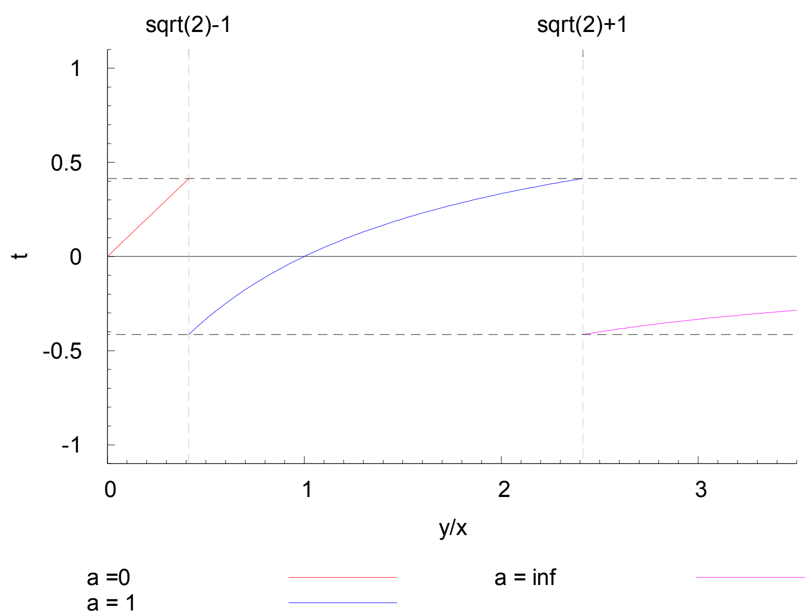
$$r=\sqrt{(2)}-1,\ \ g1=\sqrt{(2)}-1,\ \ g2=\sqrt{(2)}+1,\ \ k=1$$

So we've reduced the range of t from $-1\le t\le 1$ in our 2 range scheme to

$-.414 \leq t \leq .414$ in our 3 range scheme and the equations become:

$$ATN(s) = ATN(a(s)) + ATN(t(s)), \quad t(s) = \frac{s-a}{1+as}, \quad a(s) = \left\{ \begin{matrix} 0 & 0 \leq s \leq \sqrt{(2)}-1 \\ 1 & \sqrt{(2)}-1 \leq s \leq \sqrt{(2)}+1 \\ inf & \sqrt{(2)}+1 \leq s \end{matrix} \right\}$$

Giving:




To give us our final algorithm:

$$ATN\left(\frac{y}{x}\right) = \left\{ \begin{matrix} atn\left(\frac{y}{x}\right) & 0 \leq \frac{y}{x} \leq \sqrt{(2)}-1 \\ atn\left(\frac{\frac{y}{x}-1}{1+\frac{y}{x}}\right) & \sqrt{(2)}-1 \leq \frac{y}{x} \leq \sqrt{(2)}+1 \\ \pi/2 + atn\left(-\frac{x}{y}\right) & \sqrt{(2)}+1 \leq \frac{y}{x} \end{matrix} \right\}$$

slapping on the sign of the original argument to give our final answer of course.

Which only requires a $13^{th}$ order polynomial (7 terms)

Power Operator:

If power is 0, then return 1. Because anything raised to the $0^{th}$ power is one, including 0.
If number is 0, then return 0

If we're in the situation where we're called to compute $-b^x$ , then only integer powers of x are allowed (unless we can return complex numbers).

Test if x is an integer power.
If it is then note if it's even or odd in order to determine the sign of the answer, otherwise just exit with an error.

Now solve the problem $+b^x$
To do this we use the identity:

$$e^y = b^x \rightarrow \log(e^y) = \log(b^x) \rightarrow y = x\log(b)$$

$$e^{x\log(b)} = b^x$$

So multiply x by $\log(b)$ , then call EXP
If original number ***b*** was negative, and x was an integer power, if x was odd then make answer negative.

SQR:

Square root:

$$\sqrt{x} = x^{\frac{1}{2}}$$

Send $x^{\frac{1}{2}}$ to the Power Routine