# ELECTRONOTES 217

Newsletter of the Musical Engineering Group

1016 Hanshaw Road, Ithaca, New York 14850

Volume 23, Number 217 Sept 2013


# WHAT ABOUT: SELF-FREQUENCY-MODULATION

## INTRODUCTION:

For some time we have suggested the use of Self-Frequency Modulation (SFM) as an alternative to using a bank of standard waveshapers [1]. A single continuous control clearly offered a wide range of timbres, with a range not unlike that from a sinewave to a sharp pulse. There is a certain economy evident here, with performance consequences that might go either way: you don't get traditional waveshapes (like triangle, saw, square and pulse) but you do get a continuous array.

Frequency Modulation (FM) has a proud and useful tradition in synthesized music starting with Bob Moog's "clangorous sounds" described in his classic voltage-controlled exposition [2] and then through John Chowning's digital implementations [3]. Dynamic depth, through-zero FM was clearly a magnificent tool implemented by Moog and Bode with a frequency shifter [4], in numerous digital computations, with analog circuitry [5], and with DSP processors [6]. The case of dynamic depth with exponential control was investigated [7] and an associated undesirable pitch shift was related.

Where does SFM fit in? Is it merely as a waveshaper substitute? Our VCO circuit with SFM [1] offered dynamic depth. This does involve a pitch shift. Or does it – if we can block the DC, doesn't the pitch shift go away – at least most of it? The circuit has a 1 mfd blocking capacitor. As far as I remember, we never looked at this in great detail.

## BASIC SELF-FREQUENCY-MODULATION

Fig. 1a  Full SFM

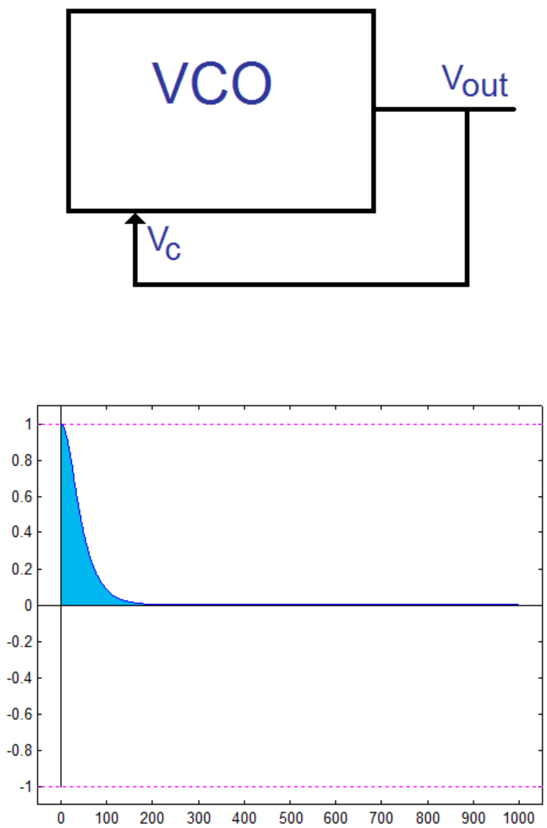
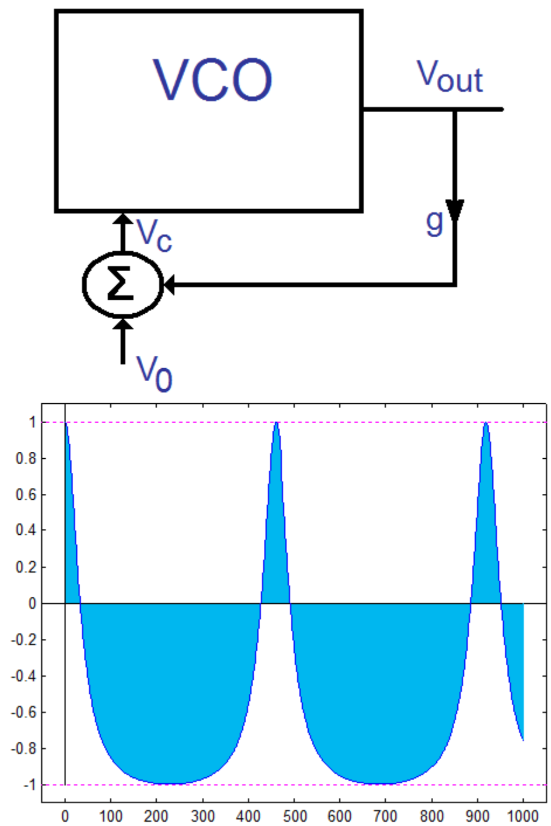


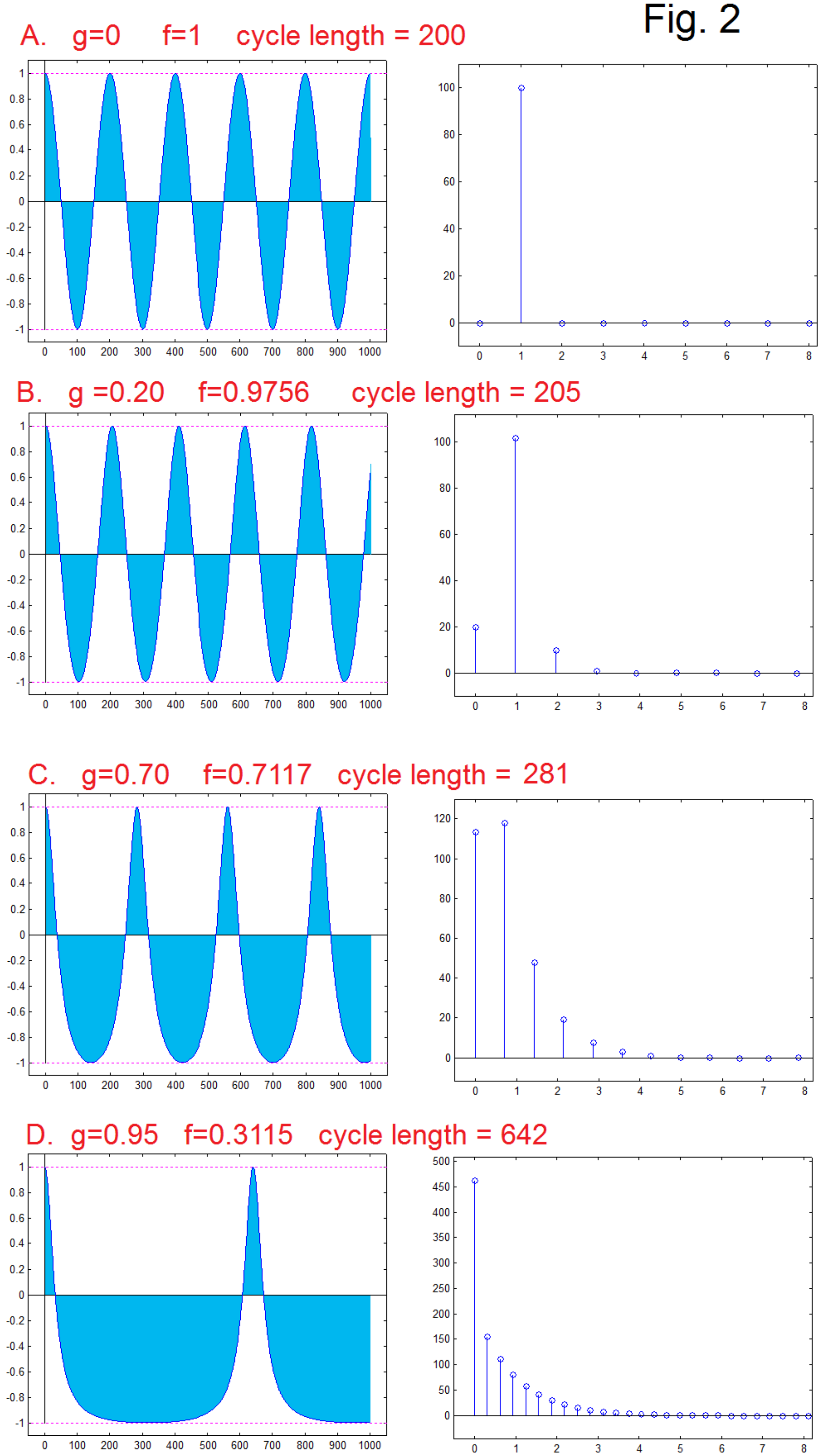

Fig. 1b Scaled SFM With Offset




The notion of SFM can be easily illustrated as having the output waveform of a Voltage-Controlled Oscillator (VCO) fed back to serve (perhaps in whole as in Fig. 1a, <u>or in the useful case</u>, just in part, as in Fig. 1b) as the control voltage input.

Keep in mind that the control voltage $V_c$ is generally thought to control the VCO frequency but it can also be thought of as controlling the rate of change of phase. Thus in Fig. 1a, where $V_c$ is connected directly to the output, $V_{out}$, as soon as $V_{out}$ hits zero, the VCO just stalls. See graph in lower left. Here we are assuming a linear VCO. Not very useful. On the other hand, as in Fig. 1b, we can choose an offset $V_o$ and a scaling of $V_{out}$ such that the sum, $V_c$ can never reach zero. Here for our example $V_o$=1 and g=0.9, so with $V_{out}$ restricted to ±1, $V_c$ can never be smaller than 0.1. [We note here that an exponential VCO would not stall.] The waveshape in the lower right plot is typical of SFM as we know it. Clearly we are a long way from a sinewave, and there is an enormous DC component is this case.

Fig. 2


A. g=0 f=1 cycle length = 200
B. g =0.20 f=0.9756 cycle length = 205
C. g=0.70 f=0.7117 cycle length = 281
D. g=0.95 f=0.3115 cycle length = 642

Fig, 2 shows more examples. See also Appendix A for graphs of the frequency and of the DC level for a continuum of g. The full Matlab code used to simulate this SFM is presented at the end of this report. The methodology is:

(1) Assume we start with a cosine. That is, the phase starts at zero and the first sample of the waveform is 1.

(2) Calculate the current value of the derivative of the phase as being proportional to the control voltage, which is the offset $V_o$ plus a scaled (by g) value of the previous value of $V_{out}$.

$$dp/dt = \pi [ V_o + gV_{out}(n-1) ] \quad (1)$$

(3) Increment the phase by multiplying by $\Delta t = 1/200$

$$p(n) = p(n-1) + (dp/dt) \Delta t \quad (2)$$

(4) Finally, take the cosine

$$V_{out}(n) = \cos[ p(n) ] \quad (3)$$

Indeed we find that as g increases, but is still held short of 1, the waveform changes from a cosine (g=0) and becomes more pointed on top (broader on the bottom). This is because as $V_{out}$ is positive, and dp/dt gets larger, the cosine speeds along, but slows greatly as $V_{out}$ goes negative. In consequence, the period gets longer, and the frequency drops below 1. Shown also in Fig. 2 are the FFTs (all FFTs here show magnitudes) , all scaled to the frequency of 1 being the original cosine. We see that while the waveform becomes "non-sinusoidal", the harmonics are those of the lower pitch. When used as a simple waveshape modifier (not dynamically through the time) we easily correct for the pitch shift by turning up the control voltage.

We do see the DC accumulating in the negative portions of the waveforms, and large DC components in the FFTs. Is this of concern? Of course, as this DC is fed back, it IS the cause of the pitch shift – if this shift is a concern, as it will be in dynamic depth. A second concern is whether or not the DC is audible. Of course, DC in the waveform is not audible, except as it may be fed through an enveloping or similar amplitude manipulation. We kind of looked at this in a related study last issue [8], which involved PWM. We kind of decided that it was not likely an issue. Indeed, looking at Fig. 2D we note significant similarity between the peaking due to SFM and a narrow pulse achieved with PWM. Likely the same results would be found here. Accordingly, we will address the issue of the pitch shift. First, let's revisit the SFM of the ENS-76, VCO Option 3, an actual circuit.

## REVISITING VCO OPTION 3

Without hesitation I would name VCO Option 3 [1] as the one circuit for which I have had the most questions over the last decade or so. I can't say that this is my favorite VCO. But on the other hand, I never had much of an opportunity to develop any favorites – it was always a design/build/test/write-up mode – and then on to something else. So here I rely, as I often do, on what I originally wrote, as tempered through later questions and occasional reflection.

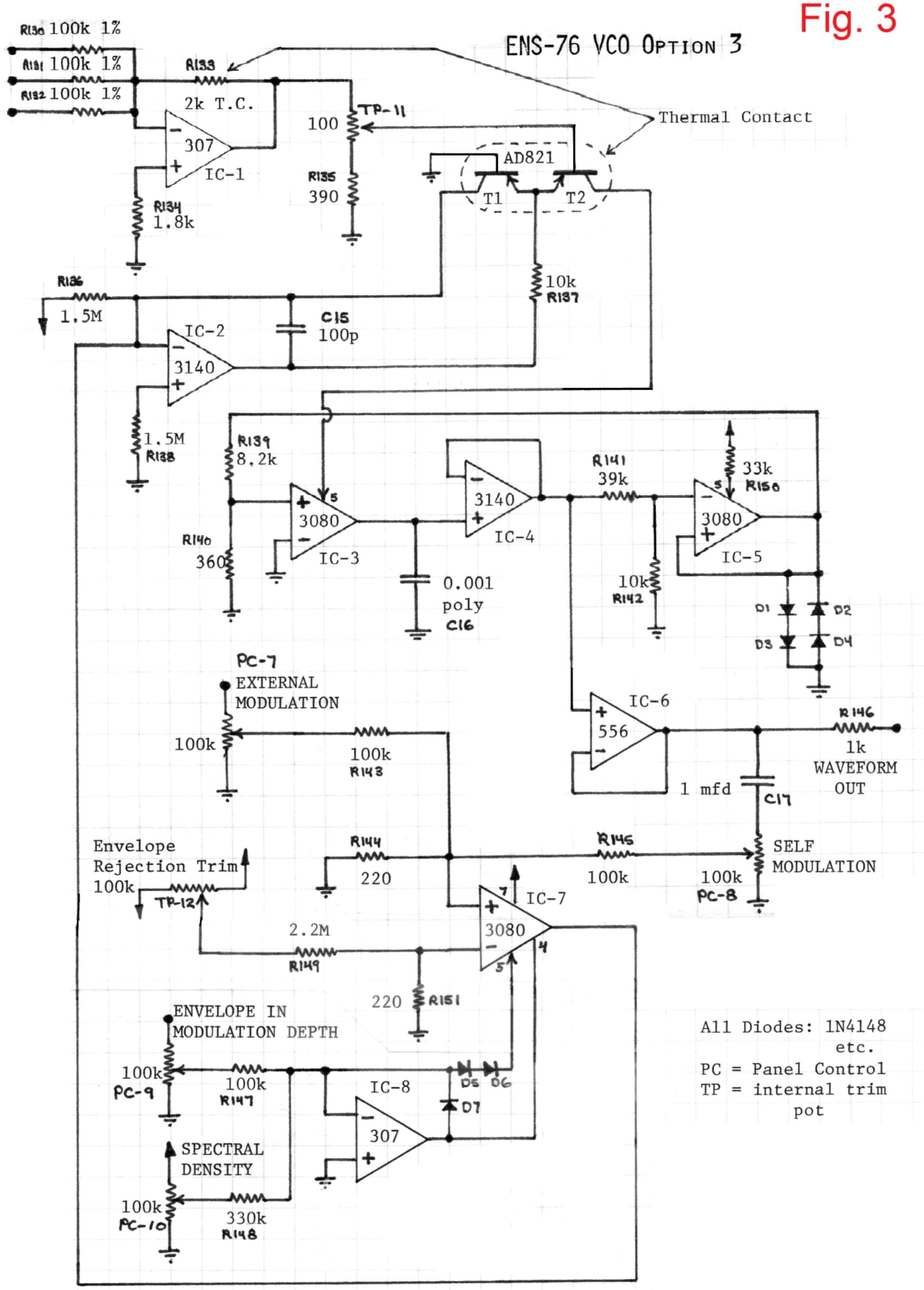
Fig. 3 is the figure as it appeared as Fig. 14 in EN#75. It is an exponential VCO with a very familiar "triangle core" and no waveshapers save the SFM feature. Not even a triangle to sine was used. Note well the inclusion of a large DC blocking capacitor (C-17) in the SFM path. This was included specifically because we also included a dynamic depth capability through IC-7. I have no recollection of studying in detail if this was highly effective in removing the pitch shift such as that suggested in Fig. 2. However, I do not recall the problem <u>not</u> being solved, or at least minimized. Here I rely on Fig. 15 of the original description, and no pitch shift was noted. Clearly this was a sketch of what was seen on the scope, and it is possible (likely?) that subtle transients could be involved. That is, small amounts of a DC or very low frequency control might come through. In such cases, we would need to ask if they were audible. And sometimes, these transients enhance a synthesized musical sound, as acoustical musical instruments generally have start-up transients.

## HIGH-PASSING THE FEEDBACK IN THE SIMULATION

Against the background of the simulations of Fig. 2 and the revisit of Fig. 3, we would like to show that if we block the DC in the simulation, we get rid of the DC components and the pitch shifting. This is tricky to think about. The harmonic generator seems first of all to be due to the peaking of the top of the waveform and the broadening of the bottom. Is this asymmetry due to the DC? Actually, the asymmetry would still be there without the DC, much as PWM has the same audible spectrum even when we removed the DC [8]. The exact waveforms will be different with different DC blocking schemes (high-pass filters) but the general ideas are the same.

This is a discrete simulation of an analog idea and we are not attempting to simulate the capacitor blocking. Thus we use a digital (discrete time) high-pass, and the simplest is just a difference between consecutive outputs. Even if you don't understand digital filters, you can understand that a DC component is the same in consecutive samples, so subtracting them removes the DC. This is what I tried first, and I <u>thought</u> it did not work. Was I really going to have to use all that digital filter theory I was teaching not that long ago!

Fig. 3


ENS-76 VCO Option 3
R130 100k 1%
R131 100k 1%
R132 100k 1%
R133
2k T.C.
307
IC-1
R134
1.8k
100
TP-11
R135
390
Thermal Contact
AD821
T1
T2
R136
1.5M
IC-2
3140
C15
100p
10k
R137
1.5M
R138
R139
8.2k
R140
360
3080
IC-3
3140
IC-4
0.001
poly
C16
R141
39k
33k
R150
3080
IC-5
10k
R142
D1
D2
D3
D4
PC-7
EXTERNAL
MODULATION
100k
100k
R143
IC-6
556
R146
1k
WAVEFORM
OUT
1 mfd
C17
Envelope
Rejection Trim
100k
TP-12
R144
220
R145
100k
100k
PC-8
SELF
MODULATION
IC-7
3080
2.2M
R149
220
R151
ENVELOPE IN
MODULATION DEPTH
100k
PC-9
100k
R147
IC-8
D5
D6
D7
307
SPECTRAL
DENSITY
100k
PC-10
330k
R148
All Diodes: 1N4148
etc.
PC = Panel Control
TP = internal trim
pot

No – because there was an indication that it was working. The DC was gone. The problem was that it looked an awful lot like no feedback at all. Was this really a "gotcha"? Well the key is the fact that there was a small indication of feedback. Then it struck me that the problem was that I was still stuck on the idea that the gain should be between 0 and some value stopping short of 1. Here I had a simple (simplest) high pass consisting of a zero at z=1 in the z-plane. For low, but non-zero frequencies, the gain of the filter itself was still tiny. The frequency of interest was effectively 1 for a sampling frequency of 200. In an analog context, we would worry about this extreme attenuation. In a simulator, you just increase the gain.

Thus step (2) of the simulation became:

$$dp/dt = \pi \{ Vo + g [V_{out}(n-1) - V_{out}(n-2) ] \} \qquad (4)$$

(see Fig. 4) and g can now exceed 1. In fact, here results were stable through g=40.

Fig. 4 High-Pass in Feedback

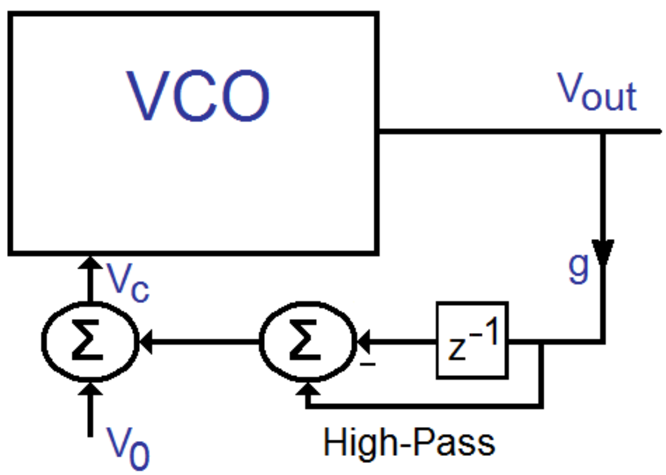
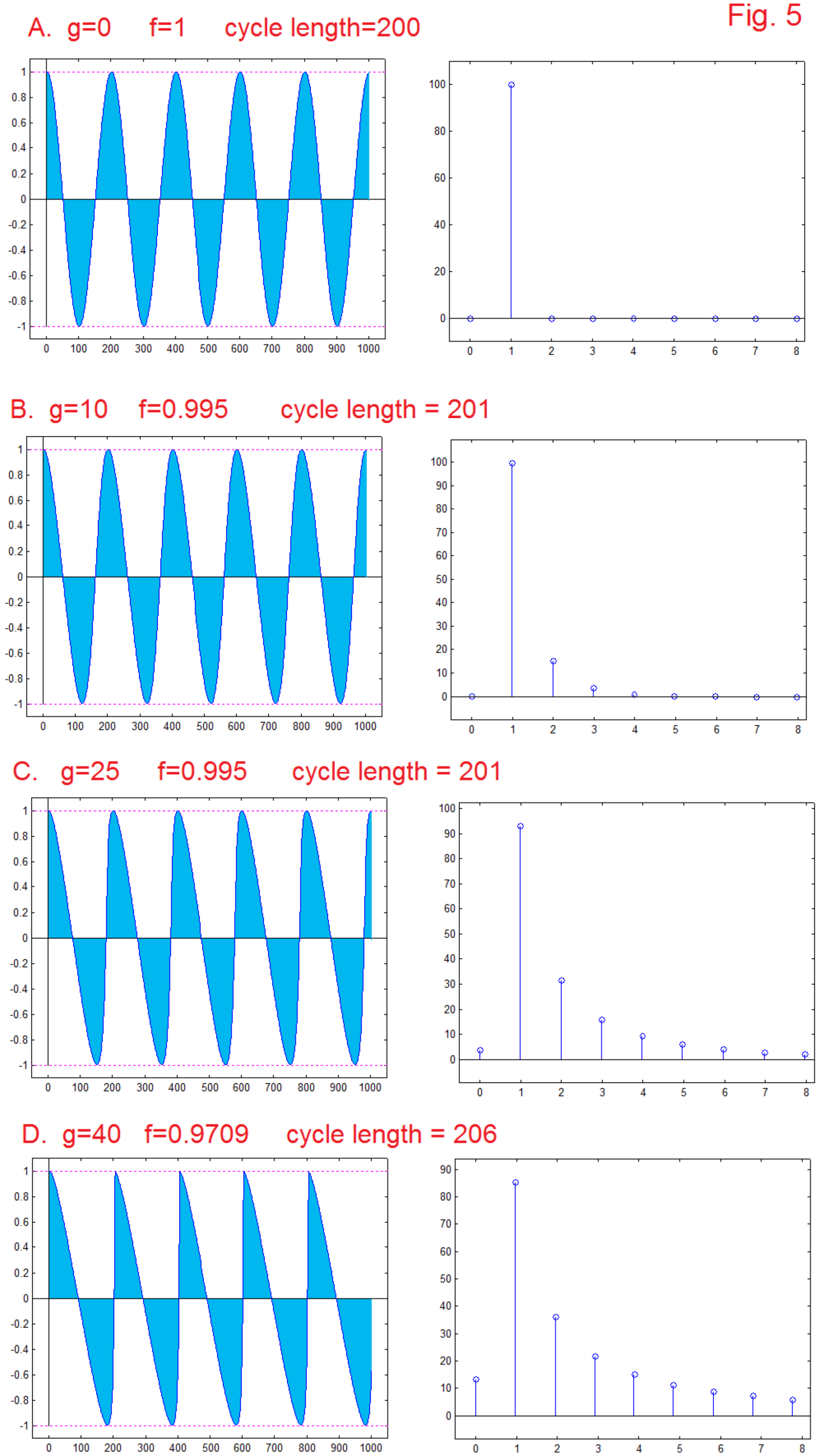



Fig. 5 shows four examples starting with g=0 in A, to g=40 in D. The results in Fig. 5 are different from Fig. 2 is several ways. Most importantly, we see good success getting the DC and resulting pitch shift blocked. For g=10 and g=25, B and C, we see very little pitch shift (the one listed may be mainly a numerical effect). In addition, the waveshape is different. Indeed it is more like a sawtooth mimic than a pulse mimic. Different DC blocking schemes likely give still other shapes. Note that the FFT shows harmonics that are multiples of the original pitch which is now very close to 1 in all the examples (not getting smaller as in Fig. 2). In some sense, this shows that dynamic depth can be used if DC blocking occurs.

Fig. 5


A. g=0 f=1 cycle length=200
B. g=10 f=0.995 cycle length = 201
C. g=25 f=0.995 cycle length = 201
D. g=40 f=0.9709 cycle length = 206

## DISCUSSION

Having just declared victory, we nonetheless feel obliged to discuss this just a bit further. In particular, how hard would it be to obtain a closed-form expression for $V_{out}(t)$ rather than the simulation $V_{out}(n)$? We can write the equations, but can we solve them?

We should perhaps begin by noting that there is nothing remarkable about self-modulation. Fig. 6 shows a situation where we have added a second VCO to our scheme, assumed identical and identically initialized. The first VCO is self-modulated and the second one is driven by the first – obviously. The first has feedback, and the second does not.

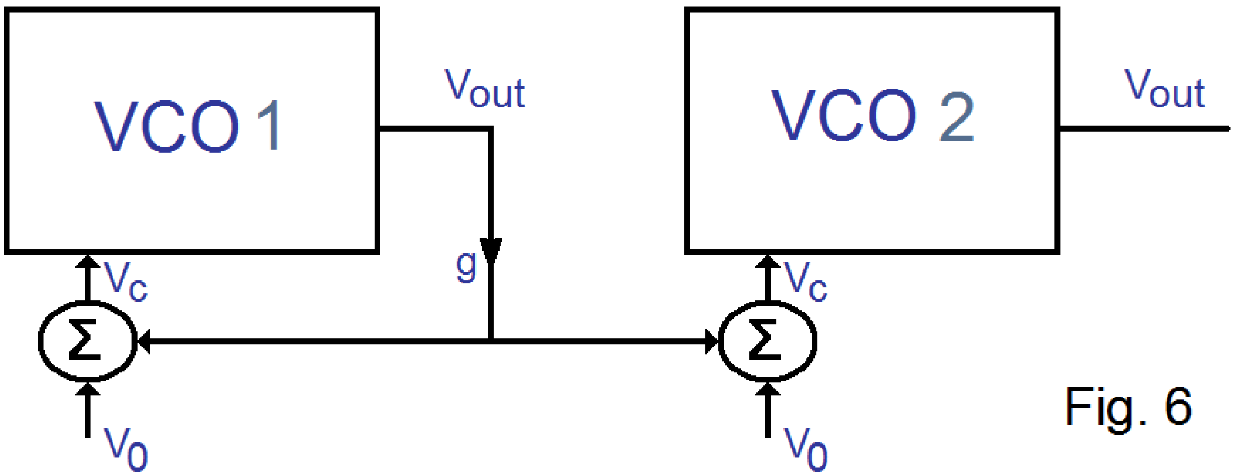



Fig. 6

Here we arrive at a point about feedback that is often overlooked or which is unmentioned. In Fig. 6, both oscillators are driven by the exact same control voltage, the essential part of interest here comes from the output of the first VCO. Does the first VCO "recognize" in some sense, that the drive is its own output? Of course not. So as far as self-modulation is concerned, the first VCO is just responding to <u>some voltage</u> that happens to be from its own output. It does not "know" this or respond in any special manner to its own product.

Thus, we have from our simulations some idea what $V_{out}$ is – a periodic waveform, a series of harmonics. We are accustomed to writing down FM equations in terms of a carrier frequency (as a sinusoid) and a program signal (also a sinusoid). Here we are talking about a program signal that has multiple sinusoidal components. Very familiar after all. Further, another often made point about music synthesis with FM is that we are modulating, and then using the modulated spectrum as a musical signal, <u>but not demodulating</u>.

A fundamental question here is whether or not the VCO is “linear”. Note however that we very often use the term “linear” to address different issues. For purposes of discussion, let’s assume the VCO is linear in the sense that the output frequency is a linear function of the control voltage. Thus:

$$f_{out} = K\,V_c \qquad (5)$$

That’s clear. We equivalently suggest that this means the output waveform of the VCO is:

$$V_{out} = \cos(2\pi f_{out} + \varphi) = \cos(2\pi K V_c + \varphi) \qquad (6)$$

At this point, we want to look at “linearity” in the alternative formal sense of whether or not superposition applies. Thus we want to consider $V_c$ to have two components $V_{c1}$ and $V_{c2}$.

$$V_c = V_{c1} + V_{c2} \qquad (7)$$

For example, suppose $V_{c1} = 2$, $V_{c2} = 3.2$, and K=100 (Hz/volt). Then using equation (5) with superposition linearity, we would have 200 Hz corresponding to 2 volts and 320 Hz corresponding to 3 volts, and $V_c = 5.2$ would correspond to 520 Hz. This is because we have chosen frequency as the output of the system being examined for linearity (the VCO).

[ Another way of stating this sense of linearity is that FM radio works. Music or other audio signals carried as ordinary radio programs have multiple components and these do not become a hopeless jumble in FM. We demodulate just fine. FM is thus seen as more of a temporary mapping or coding. We recover the program. So much for equation (5) – what about equation (6)? ]

In the view of observing the spectrum instead of the recovered signal, we would suppose a control voltage of 2 volts would produce a signal $V_{out}$ at 200 Hz. A control voltage of 3.2 volts would produce a signal $V_{out}$ at 320 Hz. But as noted, a control voltage of 5.2 volts does not produce two frequencies (at 200 Hz and at 320 Hz) but one signal at 520 Hz. This test of linearity (the output of the VCO is a voltage signal, not a frequency) fails. So the way in which a VCO handles a multi-component control signal (as linear or not) depends on what you consider to be the VCO output (the frequency/spectrum or the time signal).

We have at times found it useful to consider a VCO as an integrator. That is, the VCO has not so much a frequency, but an “instantaneous frequency” which is the derivative of the phase with respect to time. Hence the VCO integrates to total phase. This is pretty much the view used in the simulation, equations (1) and (2).

So now we get to the point where we can write down equations, but not the solution. So let's just say it is not immediately apparent how to arrive at the closed from solution. Here is the problem:

As in equation (1), we write in the continuous case:

$$\frac{dp}{dt} = \pi[V_0 + gV_{out}(t)] \quad (8)$$

$$p(t) = \int_0^t \frac{dp}{dt}\, dt = \pi \int_0^t [V_0 + gV_{out}(t)]\ dt \quad (9)$$

$$V_{out}(t) = \cos\ [p(t)] = \pi \cos[\mathrm{V_0\, t + g} \int_0^t \mathrm{V_{out}\,(t)dt}\,] \quad (10)$$

Well. Is this set up right? If so - how do we solve it?

## CONCLUSIONS – THE LOOP

As mentioned, based on the original design and tests, and what we have shown above, we believe that the SFM technique of Option 3 is valid and worthwhile. So perhaps we should remark on why a lot of people have trouble getting the circuit to work. In this respect, it is useful to note that this circuit is a loop. Most VCO circuits have a loop, the "core" oscillator. It is usually pretty clear if this is working or not. Given this determination, the waveshapers tend to be branches off this loop. Possibly one finds all but one waveshaper working, so you concentrate on that one branch. The core oscillator provides its own test signal.

With this Option 3, with its self-modulation loop, something that fails in the loop can make the whole structure inoperative. The loop provides the linear control current, and can easily shut the core oscillator down if it fails. You need to be very careful with the connections, and be prepared to bypass sections appropriately. For example, disconnecting the output of IC-7 from IC-2 and instead running the left side of R145 directly to the (-) input of IC-2, and then starting with the pot PC-8 low should show if the oscillator runs and then if the self-modulation (manually controlled) works. That sort of thing. Over the years I have made suggestions for testing the circuit, and many have gotten it to work and replied. When I don't hear, I am never sure if they succeeded to just gave up.

## REFERENCES

[1] B. Hutchins, "The ENS-76 Home-Built Synthesizer System – Part 7, VCO Options" (see Option 3), ***Electronotes***, Vol. 9, No. 75, March 1977, pp 14-16

[2] Robert A. Moog, "Voltage-Controlled Electronic Music Modules," ***J. Audio Eng. Soc***., Vol 13, No. 3. July 1965, pp 200-205
http://moogarchives.com/aes01.htm

[ 3 ] John Chowning , "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation, " ***J. Audio Eng. Soc.***, Vol 21, No. 7, Sept 1973, pp 526-534
https://ccrma.stanford.edu/sites/default/files/user/jc/fmsynthesispaperfinal_1.pdf
http://people.ece.cornell.edu/land/courses/ece4760/Math/GCC644/FM_synth/Chowning.pdf
Also see the interesting YouTube: http://www.youtube.com/watch?v=w4g92vX1YF4

[4] Robert A. Moog and Herald Bode, "A High-Accuracy Frequency Shifter for Professional Audio Applications," ***J. Audio Eng. Soc***., Vol 20, No. 6 (1972) pp 453

[5] A through-zero FM VCO was offered as an option for the ENS-76 system published in ***Electronotes*** 75 as in [1]. A better design is offered as: B. Hutchins, A Voltage-Controlled Oscillator with Through-Zero FM Capability," ***Electronotes***, Vol. 13, No. 129, Sept 1981 available here: http://electronotes.netfirms.com/EN129.pdf and a very brief summary discussion (to give proper credit to Doug Kraul and Bob Moog) is online here:
http://electronotes.netfirms.com/TZFMNote.pdf

[6] B. Hutchins, "A Review of Frequency Modulation Synthesis Techniques for Musical Sounds," ***Electronotes***, Vol. 16, No. 171, April 1988
The excerpt from this large article relevant to a digital oscillator, and the way that many issues are (fortuitously) simultaneously solved (a real free lunch) is:
http://electronotes.netfirms.com/EN171Part.PDF

[7] B. Hutchins, "The Frequency Modulation Spectrum of an Exponential Voltage-Controlled Oscillator", ***J. Audio Eng. Soc***., Vol. 23, No. 3, April 1975, pp 200-206
http://electronotes.netfirms.com/AES3.PDF

[8] B. Hutchins, "Revisiting Pulse Width Modulation," ***Electronotes***, Vol. 23, No. 216, July 2013
http://electronotes.netfirms.com/EN216.pdf

## APPENDIX A

Here we show frequency and DC for a range of g=0 to g=0.98. As g gets even closer to 1, the period exceeds the range of n examined by the program. This is enough anyway.

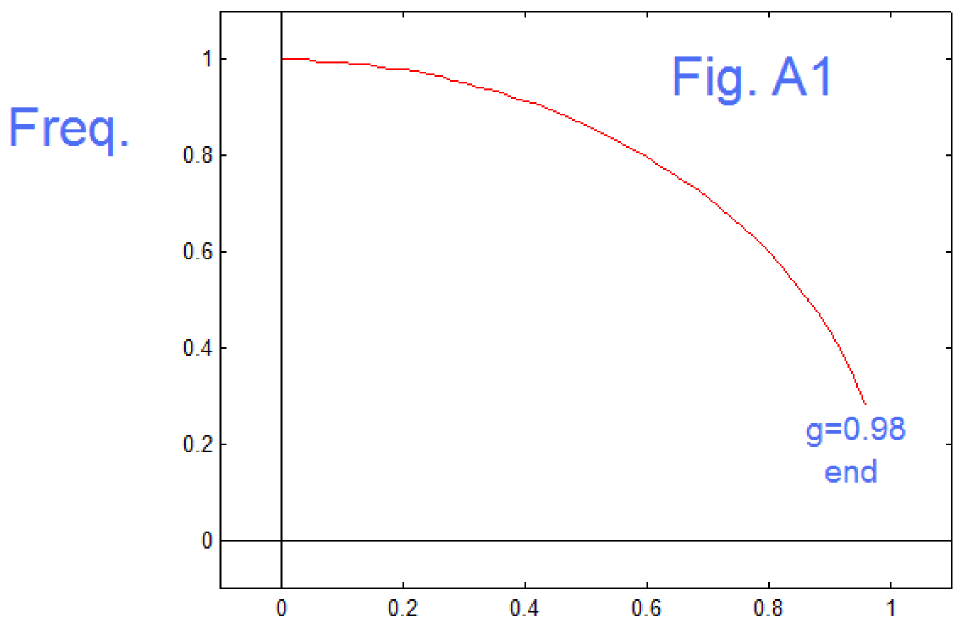



Fig. A1

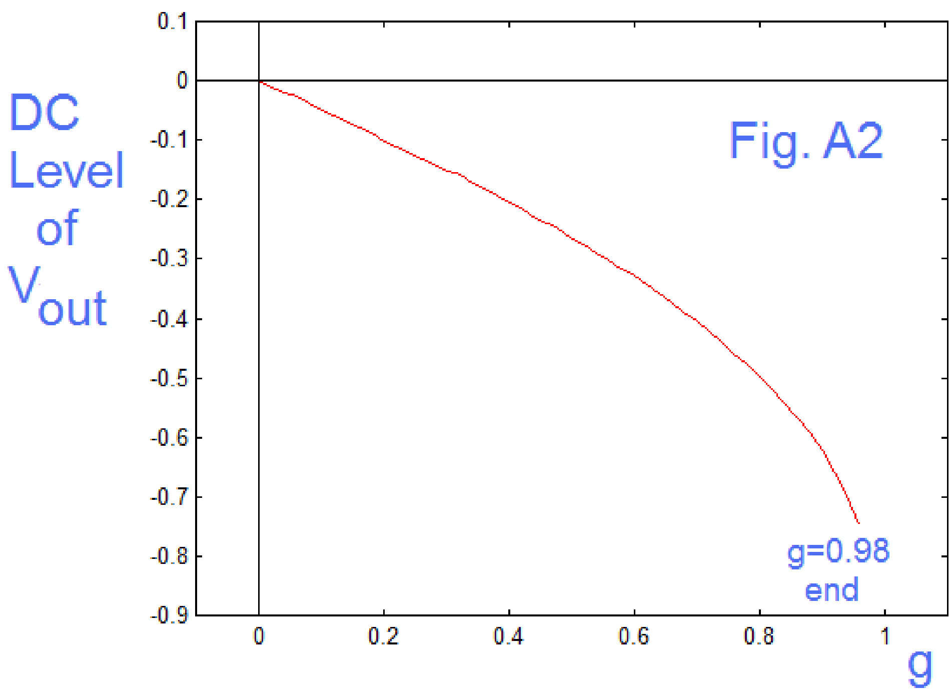



Fig. A2

## APPENDIX B – MATLAB PROGRAMS
## APPENDIX B1 FOR FIG 2

```
% sfm3.m
%   Fig. 2
clear
t=0
delt=0.01
vout(1)=1
vout(2)=2
p=0
vo=1

g=0.95
nmax=1000

for n=2:nmax
   dpdt=pi*(vo+g*(vout(n-1)));   % original - no highpass
   p=p+dpdt*delt;
   vout(n)=cos(p);
end

figure(1)

plot([-100 1100],[0 0],'k')
hold on
plot([-100 1100],[-1 -1],'m:')
plot([-100 1100],[1 1],'m:')
plot([0 0],[-1 10],'k')
plot([1:nmax] ,vout)

axis([-50 1050 -1.1 1.1])
hold off
figure(1)


maxvout=max(vout)
go=0
```

```
for k=10:1000
  if vout(k)>0.998
    if go==0
    cyclen=k+1
    go=1
    end
  end
end

cyclen
f=200/cyclen


cycle=vout(1:cyclen);
dc=sum(cycle)/cyclen
CYCLE=fft(cycle);

figure(2)
stem([0:cyclen-1]*(200/cyclen),abs(CYCLE))
hold on
plot([-100 1000],[0 0],'k')
af=max(abs(CYCLE))
axis([-0.5, 8.2 -0.1*af, 1.1*af])
hold off
figure(2)
```

## APPENDIX B2 FOR FIG. 5

```
% sfm5.m
% Fig. 5
clear
t=0
delt=0.01
vout(1)=1
vout(2)=1
p=0
vo=1

g=25
nmax=1000

for n=3:nmax
   dpdt=pi*(vo+g*(vout(n-1)-vout(n-2)));      % here is the high-pass
   p=p+dpdt*delt;
   vout(n)=cos(p);
end

figure(1)

plot([-100 1100],[0 0],'k')
hold on
plot([-100 1100],[-1 -1],'m:')
plot([-100 1100],[1 1],'m:')
plot([0 0],[-1 10],'k')
plot([1:nmax] ,vout)

axis([-50 1050 -1.1 1.1])
hold off
figure(1)


maxvout=max(vout)
go=0
```

```
for k=10:1000
  if vout(k)>0.998
    if go==0
    cyclen=k
    go=1
    end
  end
end

cyclen
f=200/cyclen


cycle=vout(1:cyclen);
dc=sum(cycle)/cyclen
CYCLE=fft(cycle);

figure(2)
stem([0:cyclen-1]*(200/cyclen),abs(CYCLE))
hold on
plot([-100 1000],[0 0],'k')
af=max(abs(CYCLE))
axis([-0.5, 8.2 -0.1*af, 1.1*af])
hold off
figure(2)
```