A SYSTEM FOR GATING SEGMENTS OF TAPED SPEECH WITH GREAT PRECISION

by A.G. Wempe

The system was first published in the report of December 1975 of
the Spring Symposium of the Dutch Association for Phonetic Sciences
held at Utrecht on 12 June 1975.

1. GENERAL

The system essentially consists of a stereo tape recorder, a
control unit and a storage oscilloscope.
The tape fragment carrying, on track I, the segment to be isolated
is searched by manual operation of the recorder and subsequently
'underlined' by recording a pilot signal on track II. This signal
is generated by the control unit as long as a push button is held
depressed. The exact length of the pilot signal is immaterial so
long as the particular segment is encompassed by it. When the
control unit is switched to START the recorder starts to rewind and
continues to do so until the complete pilot signal has passed (the
tape remains pressed against the heads in REWIND mode). After this
the control unit switches the recorder to PLAY mode, the recorded
pilot signal serving as a time base generator to control a digital
timing device.
A particular part ('segment') of the 'underlined' fragment can now
be gated by defining it in terms of time and setting the timing
device accordingly, using two sets of thumbwheel switches. As soon
as the gating is finished the recorder rewinds until the complete
pilot signal has passed again, after which the recorder stops.
If the start button is depressed again a new cycle is initiated.

2. TAPE RECORDER

The tape recorder we use in the system is a REVOX type A77, with remote control facility for all modes of tape transport and therefore not needing electrical modifications. The only required mechanical modification is the fitting of a small lever preventing the tape from being lifted off the heads during REWIND mode.
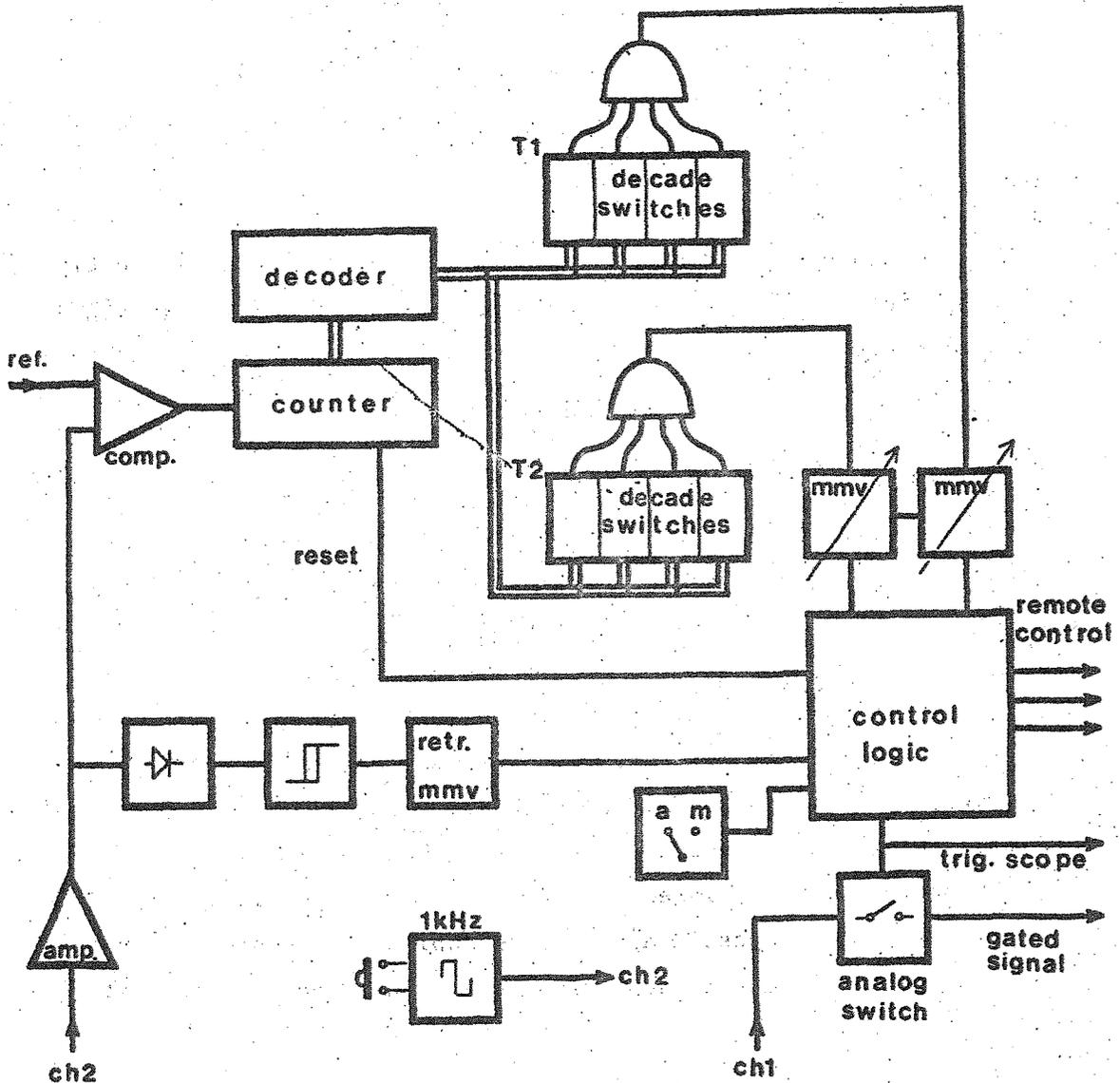

FIG. 1

3. CONTROL UNIT (see fig. 1)

The functions of the control unit are:

1. Generating the pilot signal;
2. Gating the segment in correspondance with the positions of the thumbwheel switches;
3. Controlling the tape recorder.

In addition the control unit triggers the storage oscilloscope at the beginning of the segment to be gated.

To describe these functions in some detail we start at the moment when the recorder, being in the PLAY mode, has reached the beginning of the pilot signal on track II.

The pilot signal after amplification drives an analog comparator which produces a square wave signal with a period coincidental with that of the pilot signal. Its amplitude and slope are so chosen as to ensure correct triggering of the digital counter. Every time when the pilot signal exceeds a fixed reference level the number formed by the counter is incremented by one. Thus the counter counts the periods of the passing pilot signal. Its frequency is 1000 Hz, hence the counter has a resolution of 1 msec.

The output of the counter is converted by the BCD/decimal decoder into a decimal presentation and fed to two sets of four decade switches (thumbwheel switches), one set defining T_1, the other T_2. When the BCD number formed by the counter equals T_1, the corresponding monostable multivibrator (MMV) is started. The MMVs are adjustable from about 20 μsec to 1 msec, permitting fine control of T_1 and T_2.

As soon as the MMV under consideration goes low an analog switch is closed by the control logic causing the signal on track I to be transmitted to the oscilloscope. The digital control voltage of the analog switch is used for triggering the oscilloscope so that it starts to write the signal. At this moment the counter is reset and starts to count again from zero.

Up to this moment the control logic has only been open to the
command coming from the MMV of the T_1 circuit.

From now on it is ready to receive a pulse from the T_2 circuit.
Thsi MMV is started when the counter reached the number corres-
ponding to T_2. When it goes low the track I signal is blocked
by the analog switch, the recorder starts rewinding and the
counter is reset to zero. The reset signal remains operative,
inhibiting the counter during the rewinding of the tape.

Since the tape is not lifted off the heads during rewinding the
control unit will receive the pilot signal again. The signal is
rectified, shaped by a schmitt trigger and fed to a retriggerable
MMV.

For the duration of the pilot signal the output of the retriggerable
MMV remains high, because its time constant is always greater than
any period of the pilot signal.

When the pilot signal has ceased (i.e. at the beginning of the
pilot signal on the tape) the retriggerable MMV is allowed to go
low and the control logic issues a stop command. A new START command
will not be obeyed until approximately 1.5 seconds later, an inter-
val sufficient to allow the tape to stop. If the recorder is started
again the reset signal for the counter is cancelled. The control
logic can now again admit the signal from the MMV of the T_1 circuit.
The distance between the new starting point and the beginning of
the pilot signal is sufficiently long to allow initial irregularities
such as poor contact between tape and head, excessive flutter, to
be over when the pilot signal is reached, so that its beginning is
well-defined.

(If T_1 and T_2 are inadvertedly so chosen that their sum exceeds
the number of periods of the pilot signal, the rewinding starts at
the end of the pilot signal).

4. PRECISION

It will be clear that the reliability and the precision of the system

are first of all dependent on the proper clocking of the counter.
The moments of the triggering of the counter depend, for one thing,
on the amplitude of the pilot signal.
Since this is produced by the recorder it is subject to considerable
variations in amplitude caused by drop-outs.
The dependence of the triggering on the amplitude diminishes as
the rise time of the pilot signal is diminished and as the reference
voltage level of the comparator is set closer to zero. As for the
latter, there is a limit caused by tape noise. A short rise time
can be achieved by choosing a suitable wave form or a high frequency
of the pilot signal. However, the higher the frequency the greater
the effect of drop-outs; besides if the frequency becomes too high,
the recorder cannot deal with the signal in the REWIND mode. Obvious-
ly a square wave signal should solve the problem.
Now the fundamental frequency of the square wave can be made 1 kHz
causing one step of the counter to equal 1 msec.
The rise time is now determined by the frequency range of the re-
corder.
A 1 kHz square wave signal played back by the tape recorder will
have a shape as depicted in fig. 2. It will be clear that when a
drop-out occurs the slope steepness diminishes. Fig. 3 shows the
area around zero magnified. It appears that a fixed trigger level
is reached later when there is a drop-out. Summing up the preci-
sion now depends on the steepness of the pilot signal at the out-
put of the recorder, the signal-to-noise ratio of the signal and
the loss of amplitude as a result of the drop-outs.
When using a high quality recorder and high quality tape it is
possible to attain without any difficulty a slope steepness of
1/25 of the maximum amplitude $(V_{o\ max})$ per μsec and a S/N-ratio
of 55 dB.
In order to get an impression of the degree of inaccuracy we
assume a loss of amplitude of 20 dB as a result of a drop-out
of exeptional proportion and a trigger level of the comparator
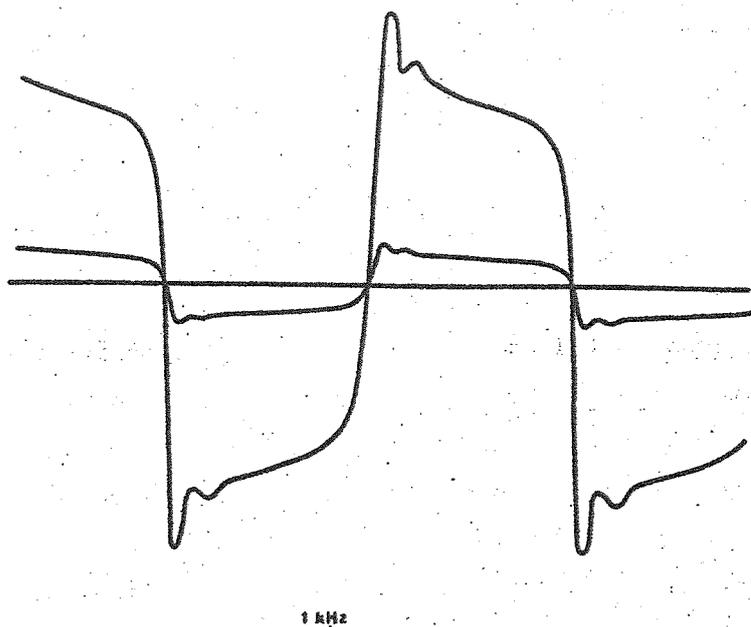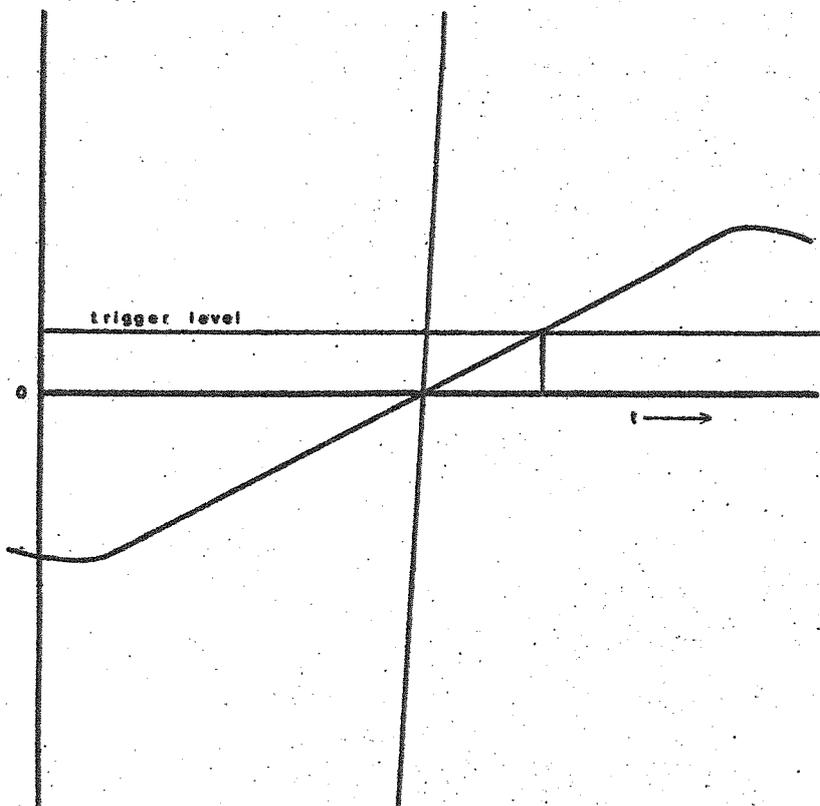of, say, 30 dB below $V_{o\ max}$, a value well above the noise level.

1 kHz

FIG.2

trigger level

0

t ⟶

FIG.3

The steepness of the slope during this drop-out is

$1/10 \times 1/25 \times V_{o\ max} = 1/250 \times V_{o\ max}$ per μsec.

The trigger level of -30 dB approximates $1/32 \times V_{o\ max}$, so this level is reached after $250/32\ \mu sec = 7.8\ \mu sec$.

In the absence of a drop-out the trigger level is reached after $25/32\ \mu sec = 0.78\ \mu sec$. Therefore the inaccuracy is about 7 μsec.

When fine controlling T_1 or T_2 an additional inaccuracy results from speed variations of the tape. At a speed variation of 0.2% (which is greater than the figure for wow and flutter of the REVOX recorder) the maximum inaccuracy that can occur is 0.2% of 1 msec = 2 μsec.

Theoretically a further inaccuracy is caused by residual elongation resulting from (repetetive) playing, rewinding and stopping the tape. (Fluctuations in elongation as a cause of wow and flutter need not concern us, because they are covered by the 0.2% mentioned above).

If the residual elongation of 0.1% quoted for the tape we use (AGFA PE 36) is valid under our working conditions the inaccuracy under consideration is 1 μsec.

Summing up, the worst case inaccuracy of the system is approximately 10 μsec, irrespective of the place and length of the segment.

The precision was tested in practice: when a short segment was registered a hundred times on the storage oscilloscope using a time base of 100 μsec per cm (i.e. a 1 msec segment) all hundred representations of the segment coincided exactly.

At the moment the gating equipment at our institute comprises the following apparatus: tape recorder REVOX A77; control unit; loudspeaker; transient recorder (electronic signal memory) BIOMATION 802; oscilloscope TEKTRONIX RM 561A; strip chart recorder HOUSTON omniscribe 5000.

The transient recorder, when scanning its memory at very low speed, enables the relatively slow chart recorder to register the segments without distortion.

The tape we generally use is AGFA PE36.

Technical data

Range between beginning of the pilot signal
and opening gate: 1 msec – 10 sec

Range of segment length: 50 μsec – 10 sec

Adjustment resolution: c. 10 μsec

Inaccuracy: < 10 μsec

S/N ratio: > 58 dB ⎫

Frequency range: 30-20.000 Hz ⎬ data of REVOX A77

Wow and flutter: ± 0.08% ⎭

Dimensions of control unit: 30 x 12.5 x 13 cm.