Balloon-Borne Radio Altimeter

NADAV LEVANON, STUDENT MEMBER, IEEE

Abstract—**This work describes an accurate pulse radar altimeter which is simple and light enough to be carried aloft by a regular sounding balloon. The altimeter uses a single superregenerative stage serving as both the receiver and the transmitter. This stage is used in a feedback system in such a way that the period between transmitted pulses is a measure of altitude. The relatively slow rate of change of altitude allows the averaging of many returns with corresponding improvement in signal-to-noise ratio.**

Accurate geometric altitude of a meteorological balloon can improve atmospheric sounding. It can also replace either temperature or pressure measurements through the use of the hydrostatic equation. Geometric altitude is essential in floating superpressure balloons, where, together with a pressure measurement, it yields the necessary reference pressure.

Analysis of the altimeter and details of the instrumentation are given, along with results of flight tests. Altitudes up to 20 km were measured in balloon flight tests over Lake Michigan. At midrange, the rms value of the random error was smaller than 0.07 percent. Height readings were available every 2 seconds during both the ascent and descent of the balloon payload. The peak power of the altimeter was one watt.

Introduction

THE radiosonde, an instrument package elevated to high altitudes in the atmosphere by a buoyant balloon, is the most widespread means of obtaining vertical profiles of atmospheric parameters. These parameters are: temperature, pressure, relative humidity, and wind velocity. The meteorologist is particularly interested in the change in these parameters with altitude. Sensors for the first three parameters are mounted in the balloon package, and their readings are transmitted to the ground. The wind is studied by tracking the balloon with optical or radio theodolites. The only parameter that is not measured directly is the altitude. However, we can allow one unmeasured parameter, because for thin layers of the atmosphere below 90 km, the following relation exists:

$$\frac{P}{P_b} = \exp\left(-\frac{g_0 M \Delta h 2}{R(T + T_b)}\right) \tag{1}$$

where

P = the pressure at the top of the layer
P_b = the pressure at the base of the layer
T = the temperature at the top of the layer
T_b = the temperature at the base of the layer
g_0 = acceleration due to gravity, at sea level
M = molecular weight of the air
R = universal gas constant
Δh = the thickness of the layer (in geopotential meters).

From (1) we see that once the pressure and temperature at a certain altitude (usually the surface) are known, the height measurement is not needed to get the vertical profile, as long as the temperature and pressure are measured at close intervals. However, at high altitudes these parameters are hard to measure accurately. For example, the barometer readings are in steps of 1 mbar, which is 0.1 percent of the total pressure at the surface, but which is 10 percent at an altitude of 30 km. The temperature readings are also less accurate at high altitude as a result of stronger solar radiation. A direct measurement of altitude at this range could improve the accuracy of the vertical profile.

The meteorological community is devoting much effort to several programs of global weather observations. Playing major roles in these programs are superpressure balloons, carrying their sensors at constant-density altitudes and floating with the winds for several months. These balloons will float most of the time over the ocean where no ground stations exist to give the necessary reference readings of surface pressure. For these balloons, a radio altimeter and a pressure measurement could supply such a reference.

The simultaneous reading of pressure and geometric altitude is most useful when we wish to study an unbalanced atmosphere. For example, the balloon in a balanced atmosphere will follow the wind along isobaric lines. Therefore, constant-pressure readings will be accompanied by constant-altitude readings, but when the geometric altitude changes separately from the pressure altitude, the wind crosses isobars, implying that a weather disturbance is growing or filling.

The above are just a few examples of the great need for a radio altimeter on board the balloon package. However, conventional radio altimeters are too heavy for balloons. It seems that just using modern circuit techniques like integrated circuitry would not be enough in reducing weight, size, and price; and that in addition, a simpler system is needed. It should be noted that the balloon offers one big advantage—its low rate of altitude changes (compared to rockets, for example). This allows plenty of time for averaging, which in turn can replace strong signals, or a good noise figure. This advantage is fully exploited in the proposed balloon-borne radio altimeter.

Manuscript received May 7, 1969. This research was supported by NASA under Contract NAS5-11542.

The author is with the Space Science and Engineering Center, University of Wisconsin, Madison, Wis. 53706.

Principle of Operation

The proposed radio altimeter uses a single superregenerative stage [1] serving as both the receiver and the transmitter. This method was used for beacons in the early days of radar. This superregenerative stage is used in a feedback system in such a way that the period between transmitted pulses is a measure of altitude. The following section describes this operation in more detail.

The superregenerative stage, when triggered by the quench pulses, produces RF pulses whose envelope area depends on the RF signal existing at the input to the stage at the beginning of each quench pulse. The operation of a superregenerative stage is described in more detail in the next section.

With no signal, the noise is the only input. When there is a signal, and when the period of the quenching pulses is equal to the delay of the return pulse, the superregenerative stage detects its own pulses (Fig. 1). As we change the period (or repetition rate) of the quench pulses, we get a maximum output from the superregenerative detector near the repetition rate,

$$f_q = \frac{c}{2h} \tag{2}$$

where

f_q = the repetition rate of the quench pulses
c = velocity of light
h = the altitude.

Varying the repetition rate around this value is equivalent to scanning the returned pulse. As we will see later, the returned pulse is close to an imperfect integration of the transmitted pulse, and, therefore, has a definite peak at the end of the transmitted pulse.

The rest of the circuitry is a servo loop designed to lock on that peak. This is an example of a "single-dimensional sinusoidal perturbation adaptive system," described in detail in the current literature [2].

The principle of such a system is shown in Fig. 2. This figure demonstrates how a sinusoidal perturbation of the parameter X, on which we have control, serves to take the derivative of a given function, $\rho(X)$, of this parameter. At the peak, the output signal at the perturbation frequency is zero, while off the peak there will be some slope, either positive or negative, and therefore an output signal, in phase or out of phase, respectively. This type of signal can be used as an error signal in a servo loop.

The block diagram of the system is shown in Fig. 3. The voltage controlled oscillator (VCO) determines the repetition rate of the quench pulses. This rate is linearly related to the voltage at the output of the integrator and is sinusoidally perturbed. In other words, the quench frequency is frequency modulated at the perturbation frequency f_p; the deviation of this FM is very small. The curve of the output of the superregenerative detector as a function of the quench frequency serves

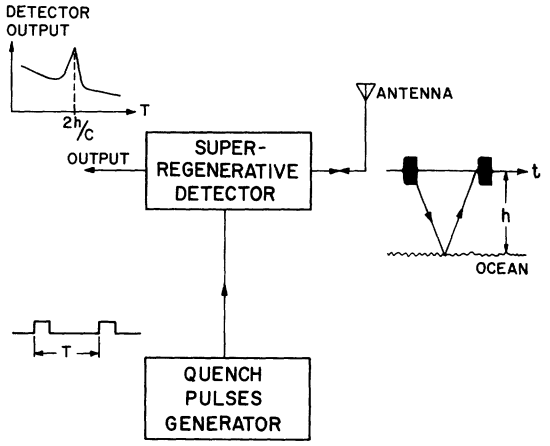

Fig. 1. The output of the superregenerative detector peaks when the period of the quench pulses equals the delay of the returned pulse.

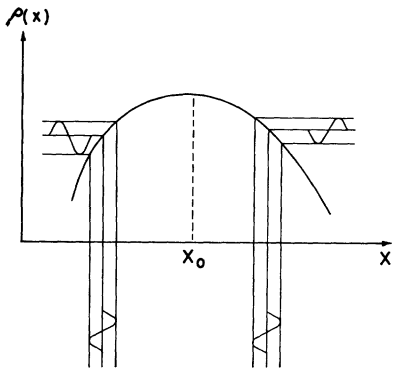

Fig. 2. The effect of sinusoidal parameter perturbations upon a single-dimensional even-function curve.

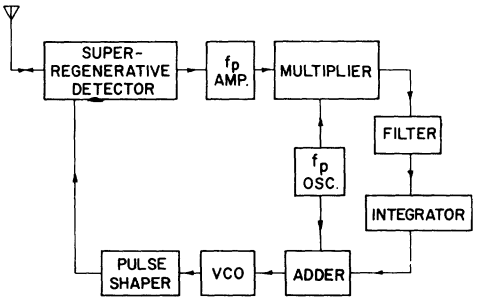

Fig. 3. Block diagram of the altimeter.

as an FM detector for the perturbation signal. When the quench frequency is too high, the detected perturbation signal is out of phase, compared to the modulating perturbation signal; when f_q is too low, the perturbation signal is in phase; and when it is at the right frequency, we get only the second harmonics of the perturbation signal.

The output of the detector is passed through a narrow-band amplifier centered around f_p and then multiplied by the modulating signal. The output of this multiplier contains a positive dc voltage when the two signals are in phase, and a negative dc voltage when they are out of phase. The integrator output is affected by

this dc voltage, and locks the VCO to that frequency which gives the peak at the output of the superregenerative detector.

The Superregenerative Detector

The fundamental action in a superregenerative circuit is centered around the growth of oscillations in an oscillator. The variation in grid voltage is shown in Fig. 4(a). Between positive pulses, the grid is at the negative potential which biases the tube beyond E_{co}, the cutoff voltage. During the positive pulses, the grid potential is raised to a point where the circuit oscillates, as indicated by projection into the shaded oscillation region. Thus, oscillation grows during the A periods and decays during the B periods.

There are two modes of operation defined for superregeneration: a linear mode [Fig. 4(b)], and a logarithmic mode [Fig. 4(c)]. The linear mode results when the positive quench period A is so short that the oscillations do not have time to build to full saturation amplitude. The logarithmic mode occurs when the A period is sufficiently long to allow oscillations to build to full amplitude before the end of the A period. In our case the circuit operates also as the transmitter, and we need as much output power as possible. Therefore, we operate in the logarithmic mode.

Before saturation is reached, the envelope of the oscillations will rise as:

$$e_c = V_0 e^{Rt/2L} \tag{3}$$

where

V_0 = the input voltage when period A begins
R = the total negative resistance of the circuit
L = the circuit inductance.

When there is no input signal, the input is the noise voltage V_n. When a signal voltage V_s is present, the voltage at the start of the oscillations will be $V_s + V_n$ and (3) becomes

$$e_c = (V_s + V_n) e^{Rt/2L}. \tag{4}$$

As can be seen in Fig. 5, a higher initial voltage would cause a shorter rise time period before saturation, and, therefore, a larger pulse area. The additional area in the signal-originated pulse is added with each additional pulse, and the total change appears at the output of the detector. The change in average voltage of the detected output ΔV, assuming linear detection, is given by [1]

$$\Delta V = f_q E_{\max} \frac{2L}{R} \ln \left(\frac{V_s + V_n}{V_n} \right). \tag{5}$$

This change in output divided by the change in the input gives the gain of the stage. Assume, for example, that the signal is equal to the noise; then the change in the input is equal to the noise, which is on the order of

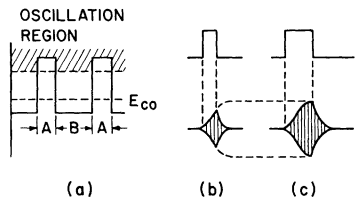

Fig. 4. Quench voltage (a) and relation to modes of oscillation (b) and (c).

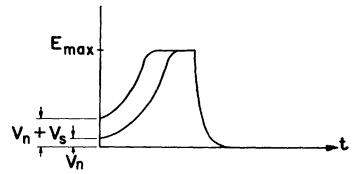

Fig. 5. The change in pulse area in the presence of signal.

a few microvolts. The change in the output, however, will be

$$\Delta V = f_q E_{\max} \frac{2L}{R} \ln 2, \tag{6}$$

which is of the order of volts. This means that a gain of a million can be achieved in a single stage.

It is for both the high receiver gain and the doubling of the single stage as a transmitter and receiver that the superregenerative stage has been chosen for a small, light, and accurate balloon-borne radio altimeter.

The Returned Pulse

The ocean surface plays the role of an imperfect integrator on the returned pulse. Let us assume, for a moment, that the three following conditions exist:

1) The average scattering cross section of the ocean is independent of the incident angle.
2) The altimeter antenna is isotropic.
3) The additional attenuation due to a longer path length of a side reflection compared to a downward reflection can be neglected.

With these three assumptions, we will follow the pulse on its way to the ocean surface and back. The first reflection occurs when the front end of the pulse reaches the point B (Fig. 6) just below the altimeter. Side returns like that from point C will start later, as a result of the additional distance, and will reach the altimeter even later because the returning pulse also has the additional distance to traverse. While the front edge of the pulse is being reflected from point C, later portions of the pulse reach point B, and start to be reflected.

In other words, the pulse can be seen as traveling on the ocean surface. When the trailing edge is being reflected from point B, the whole pulse is reflected to the altimeter, and the intensity of the return reaches a maximum. With the above three assumptions, this level will remain constant as the pulse on the ground travels away from the altimeter subpoint.

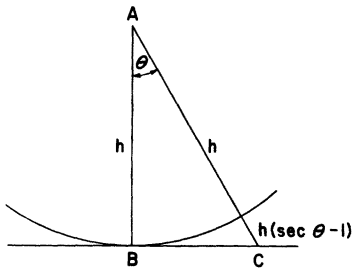

Fig. 6. The geometry of reflection.

Measurements have demonstrated [3] that with occasional exceptions, a radar return from the ground is largely due to area scatter. For this reason, no phase relations were introduced, and when we speak of the returned pulse, we are referring to the mean pulse envelope.

As we mentioned above, the maximum of the returned pulse occurs at the end of the transmitted pulse (if we ignore the basic delay). If we now remove the assumption that the antenna is isotropic, and, instead, assume that it has constant gain for $|\theta| < \theta_0$ and zero gain outside, we can calculate the minimum θ_0 that still allows the peak to occur at the end of the transmitted pulse. From Fig. 6 we see that the return from point C has additional delay τ, which is, if we assume a flat surface,

$$\tau = \frac{2h}{c}(\sec\theta_0 - 1) \tag{7}$$

where

h = the altitude
c = the velocity of light.

If the length of the transmitted pulse is shorter than τ, then the peak of the return pulse occurs at the end of the transmitted pulse, which does not change with altitude. This is called "pulse-length-limited illumination." If the length of the transmitted pulse is larger than τ, the peak will occur at τ which in turn is dependent on altitude. This is called "beam-width-limited illumination."

Our system locks on the peak of the returned pulse, and we would prefer the additional delay to be constant; therefore, we use pulse-length-limited illumination. The condition for such illumination is that

$$\sec\theta_0 \geq \frac{c\tau_0}{2h_{\min}} + 1 \tag{8}$$

where

τ_0 = the length of the pulse
$h_{\min}$ = the minimum altitude.

When we include the dependence of the average scattering cross section of the ocean on the angle of incidence, the additional attenuation due to additional distance, and the fact that the antenna pattern is not constant as a function of θ, the returned signal can be expressed as the convolution of the transmitted pulse waveform in power units with a function which includes effects of antenna pattern, ground properties, and distance.

Moore *et al.* [4] shows that for a flat earth and a system that is independent of azimuth orientation, the mean received pulse envelope in power units is given by

$$P_R(d) = \frac{\lambda^2}{8(2\pi)^2}\int_h^{cd/2} \frac{1}{r^3} P_T\left(d - \frac{2r}{c}\right) G^2(\theta)\sigma_0(\theta)\,dr, \tag{9}$$

where

r = the distance from the altimeter to the points of reflection ($r = h \sec\theta$)
h = the altitude
θ = the angle from the vertical
d = the time from the leading edge of the transmitted pulse ($d > 2h/c$)
λ = the wavelength of transmitted energy
$G(\theta)$ = the antenna gain
$\sigma_0(\theta)$ = the average scattering cross section of the ocean
$P_T(\ )$ = the envelope of the transmitted pulse in power units
$P_R(\ )$ = the envelope of the received pulse in power units
c = velocity of light.

We will include both the antenna pattern and the ground scattering pattern in one equation,

$$G^2(\theta)\sigma_0(\theta) = G^2\sigma_0\cos^6\theta = G^2\sigma_0\left(\frac{h}{r}\right)^6 \tag{10}$$

which is a good approximation in many cases.

Inserting (10) in (9), we get

$$P_R(d) \approx \frac{\lambda^2 G^2\sigma_0 h^6}{8(2\pi)^2}\int_h^{cd/2} \frac{1}{r^9} P_T\left(d - \frac{2r}{c}\right) dr. \tag{11}$$

Further approximations lead to

$$P_R\left(t + \frac{2h}{c}\right) \approx \frac{\lambda^2 G^2\sigma_0 c}{16(2\pi)^2 h^3}\int_0^t P_T(t-\tau)\cdot\left(1 - \frac{9c}{2h}\tau\right)d\tau \tag{12}$$

where t is the time from the reception of the leading edge.

In Fig. 7 a typical superregenerative pulse is used as the transmitted pulse, P_T. When this envelope is used in (12), the envelope of the received pulse is as shown in Fig 7.

For pulses of length in the order of 1 μs, the convolution integral can be approximated by a regular integral, and this will give a simpler expression for the peak of the received pulse P_2:

$$P_2 = \frac{\lambda^2 G^2\sigma_0 c}{16(2\pi)^2 h^3} P_{\max}(T - t_1). \tag{13}$$

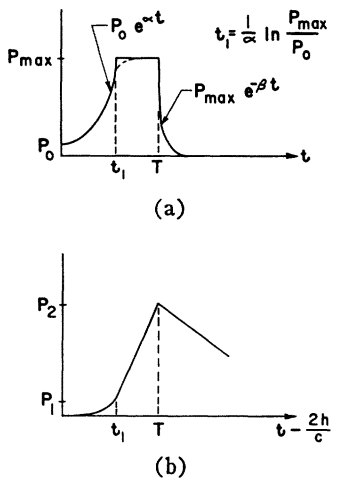

Fig. 7. The transmitted (a) and received (b) pulses.

The only constant in (13) that requires clarification is the average scattering cross section σ_0. Edison *et al.* [3] measured this value for vertical incidence at 415 MHz and found it to vary from 0.7 for dense woods, to 4 for certain city targets, and up to 50 for slightly rough water.

Differences between water and land exist also in the range of fading (defined here as the range between the level below which only 5 percent of the return power is found, and the level below which 95 percent is found). This range was as small as 3 dB for a smooth water surface and as large as 19 dB for certain nonhomogeneous target areas. The effect of this fading range could not be included without knowing the correlation of the signal fluctuations and the number of pulses averaged. The last question can be answered immediately. If the altimeter averaging period is about one second and the smallest repetition rate is 5×10^4 pps, then at least 50 000 pulses are averaged. As to the correlation of the signal fluctuations, data was unavailable for such a high repetition rate. Edison *et al.* [3] showed that at a repetition rate of 400 pps, over wooded area, consecutive returns do not resemble each other at all. From this we can conclude that at this rate the correlation coefficient is close to zero.

For the case of uncorrelated signal fluctuations, Schwartz [5] has shown that if the number of pulses averaged is larger than about 100, and the instantaneous reflected power obeys a Rayleigh probability-density function, which is typical for large targets, then the fluctuations have no effect on the detectability of the signal. Accepting Edison's results we know that we average at least 400 uncorrelated pulses. We can therefore assume that the mean returned pulse is a good approximation of the returned pulses.

Let us introduce some typical values in (13):

$P_{max} = 1$ watt
$P_0 = 4\times10^{-12}$ watt
$T - t_1 = 0.5\times10^{-6}$ second
$\sigma_0 = 1$ (for land), 20 (for water)

TABLE I
THE EXPECTED RANGE OF THE ALTIMETER

Antenna \ Surface	Land $\sigma_0=1$	Water $\sigma_0=1$
Yagi $G=3.5$	7 km	20 km
Rhombic $G=13$	17 km	46 km

$G = 3.5$ (Yagi), 13 (rhombic)
$\lambda = 0.7$ meter,

and ask what is the height h at which $P_2 = P_0$. The answer is summarized in Table I.

LINEAR ANALYSIS OF THE ALTIMETER

The sinusoidal perturbation serves to take the derivative of a curve which is actually the dependence of the superregenerative detector output on the input to the VCO. This curve includes in it the delay, intensity, and shape of the returned pulse, the linear voltage to frequency characteristic of the VCO, and the logarithmic gain of the superregenerative detector. To simplify our analysis we will assume this curve to be a parabola. The amplitude of the signal in the perturbing frequency, f_p, as it appears after the amplifier is determined by the derivative of this parabola, and is therefore a linear function of the distance from the peak of the parabola.

We are thus able to draw a linear simplified equivalent block diagram of the altimeter Fig. 8. In this block diagram c_1 is the amplitude of the perturbing signal, ρ'' is the second derivative of the parabola (a constant). A is the total gain, and it includes the gain or attenuation of

1) the RF path
2) the superregenerative circuit
3) the bandpass amplifier
4) the integrator ($=1/RC$)
5) the VCO
6) the multiplier ($=\frac{1}{2}$).

$F(s)$ is the Laplace transform of the total response of the networks in the loop, and it includes

1) the bandpass response of the f_p amplifier transferred to low pass
2) the filter response.

Of these two, the second is the dominant.

$1/s$ represents the integrator, and $n(t)$ is the noise voltage; $n(t)$ is multiplied by $1/\sqrt{2}$ to indicate that the synchronous detection of the perturbation signal eliminates half of the noise power accompanying it. The output of this loop could be either in voltage form, if it is measured at the input to the VCO, or in frequency form, if it is measured at the output of the VCO. This frequency, which is equal to the pulse repetition rate, is in essence the height information.

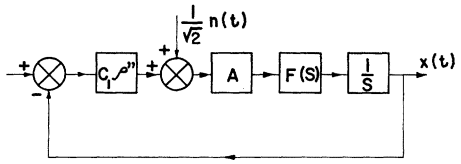

Fig. 8. A simplified equivalent block diagram of the altimeter.

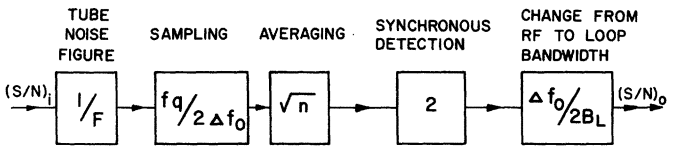

Fig. 9. Signal-to-noise ratio budget.

The loop of Fig. 8 resembles a typical phase-locked loop. We can thus use studies made on this type of loop [6], [7], in order to optimize the output error due to noise, the steady-state error, and the acquisition behavior of the loop.

The conclusions drawn from phase-locked loops led to the inclusion of a lag-lead network as the loop filter, to a choice of damping factor ξ of 0.7, and loop-noise bandwidth B_L of 1 Hz.

The primary noise source of the altimeter is the superregenerative stage. As Hall [1] showed, the effect of the sampling done in the relatively broad RF stage is to multiply the effective noise bandwidth of the rest of the receiver by twice the ratio between the RF bandwidth Δf_0 and the sampling rate f_q.

Viterbi [6] has shown that in a linear phase-locked loop the mean square error due to noise σ_x^2 is equal to the inverse of the signal-to-noise ratio at the output, $(\mathrm{SNR})_o$, i.e.,

$$(\mathrm{SNR})_o = \frac{1}{\sigma_x^2} \cdot \tag{14}$$

We can now give a crude yet meaningful answer to the following question. Given the allowed mean square error σ_x^2, what is the minimum signal-to-noise ratio per each pulse at the input of the altimeter?

To answer this question we shall use Fig. 9 in which the various processes which the signal passes, and their effect on the signal-to-noise ratio, are shown.

The first term is the $1/F$ change in signal-to-noise ratio due to the noise figure F of the RF stage. The second is the increase in noise due to sampling. The $\sqrt{u}$ improvement is due to averaging of n pulses [8].

The term 2 is a result of the synchronous detection, and the last term is the improvement in signal-to-noise ratio due to the change from the RF bandwidth Δf_0 to the loop bandwidth B_L.

We note that the number of pulses averaged, n, is given by

$$n = \frac{f_q}{B_L} \tag{15}$$

where

f_q = the repetition rate
B_L = the loop effective bandwidth.

Thus we get from Fig. 9 that the signal-to-noise ratio at the output, $(\mathrm{SNR})_o$, is related to the signal-to-noise ratio at the input, $(\mathrm{SNR})_i$, as follows:

$$(\mathrm{SNR})_o = \frac{1}{2F}\left(\frac{f_q}{B_L}\right)^{3/2}(\mathrm{SNR})_i. \tag{16}$$

Using (14) we get

$$(\mathrm{SNR})_i = 2F\sigma_x^{-2}\left(\frac{f_q}{B_L}\right)^{-3/2}. \tag{17}$$

Equation (17) is the answer to the question we have just asked. However, we will go one step further to make it more meaningful.

The noise power at the input is given by

$$N_i = KT\Delta f_0, \tag{18}$$

where

K = Boltzmann's constant
T = noise-source temperature, °K.

Inserting (18) in (17) we get for the signal power input

$$S_i = KT\Delta f_0 2F\sigma_x^{-2}\left(\frac{f_q}{B_L}\right)^{-3/2}. \tag{19}$$

It will be interesting to put numbers in this equation. We recall that σ_x^2 is the mean square error due to noise. If the total error allowed is 0.03 percent, a sensible choice of σ_x will be 0.01 percent or 10^{-4}; thus, we have for the balloon flight model:

$$KT = 5 \times 10^{-21}\ \mathrm{W/Hz}$$
$$\Delta f_0 = 5 \times 10^6\ \mathrm{Hz}$$
$$F = 8$$
$$\sigma_x = 10^{-4}$$
$$f_q = 10^5\ \mathrm{Hz}$$
$$B_L = 1\ \mathrm{Hz}$$

and the resulting signal power required at the input is $S_i = 10^{-12}$ watts, which is 7 μV over 50 ohms.

The mean square error due to noise could be further lowered if a smaller loop bandwidth B_L were chosen. The theoretical lower limit on B_L should have been determined by the rate of change of height that the altimeter has to follow. However, some practical considerations enter here. The total delay range that the altimeter can measure is considerably larger than the deviation of the returned pulse. Unless the VCO period falls within the delay range covered by the returned pulse, the loop is not closed and no acquisition is pos-

sible. To overcome this problem, a constant voltage is fed to the integrator to cause a certain sweeping rate[1] of the VCO frequency, and thus, the whole range is scanned. This sweeping rate causes a steady-state error; and in addition, it could carry the transient error beyond the hold-in range, and thus prevent locking altogether. The locking behavior of the loop in the presence of this sweeping rate is similar to phase-locked loop tracking a frequency step. For the linear model it was found [9] that the relation between the sweeping rate and the lower limit to the loop-noise bandwidth $B_{L\min}$ is given by

$$B_{L\min} = \frac{\xi + \frac{1}{4\xi}}{2\Delta T(0.6 + 2\xi)} \tag{20}$$

where ΔT is the time (in seconds) required to scan the leading edge of the returned pulse. In laboratory tests of the altimeter it was found that the practical lower limit is about three times larger than the theoretical.

Again, we shall insert practical numbers in (20). The delay range covered by the altimeter is 12 to 27 μs, and it is scanned in 30 seconds; the leading edge of the returned pulse covers 0.5 μs. From this we can conclude that $\Delta T = 1$ second. With the damping factor $\xi = 0.7$, we get for the theoretical lower limit on the loop-noise bandwidth:

$$B_{L\min} = 0.27 \text{ Hz}.$$

The practical lower limit should be 3 times larger, and this would lead us to the choice of a loop-noise bandwidth of 1 Hz.

Ambiguity

There is no reason why the altimeter cannot transmit a second pulse before the previous one is received. We shall call this subharmonic operation (Fig. 10). This mode of operation can happen by itself at any altitude above twice the minimum altitude of operation. This immediately indicates that the minimum range of altitudes required to get full coverage is equal to the minimum altitude. If the minimum altitude of operation is 2000 meters, then the range of the altimeter should be from 2000 meters to 4000 meters. Above 4000 meters the altimeter will switch to a subharmonic mode. From Fig. 11 it is clear that the first ambiguity problem starts at 6000 meters, which is 3×2000 meters, or 2×3000 meters, and the altimeter can lock on both. The pattern of switching modes as the balloon ascends, in the case of no loss of locking except at the end of range, is marked in Fig. 11. However, this pattern is not guaranteed since loss of locking might occur. In any case, the ambiguity is at least the lowest altitude (in our example

[1] It should be pointed out that some sweeping rate is inherent in the system, and some is built-in due to circuitry offsetting.

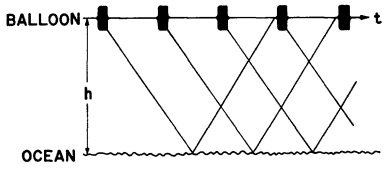

Fig. 10. Subharmonic operation.

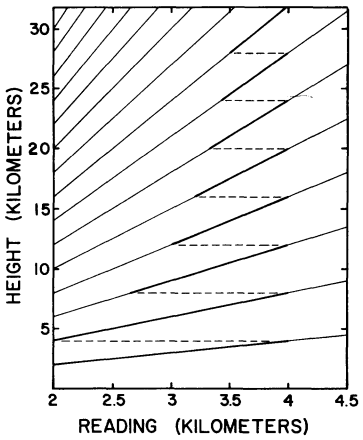

Fig. 11. Ambiguity pattern.

2000 meters), and it can be resolved by pressure reading or by the history of the ascending.

Subharmonic operation is recommended even if the range of operation is limited, as in the case of the superpressure balloons. In this case, however, it is possible to assure that only one mode will cover this limited range. Operation with subharmonics means higher quench frequency f_q and therefore better signal-to-noise ratio.

Circuitry, Laboratory Test System, and Ground Station

This section briefly covers the main features of the altimeter circuitry and the methods used for laboratory testing and reception of altimeter data.

The superregenerative stage utilizes a nuvistor triode for the 403 MHz version, and a pencil tube with integral resonator (the type used in radiosondes) for the 1680 MHz version. The narrow-band amplifier is an operational amplifier with a twin T network in the feedback loop. The multiplier is a ring demodulator. The integrator and oscillator utilize operational amplifiers. The VCO is an astable multivibrator in which the charging resistors are returned to a variable input. The pulse shaper includes a monostable stage that determines the duration of the quench pulse and an amplifier.

In addition, a bias control stage which is fed by all the supply voltages involved in the superregeneration action keeps the tube bias at the correct level in order to assure proper superregeneration. Altogether there are seven integrated circuits, one transistor, and one vacuum tube.

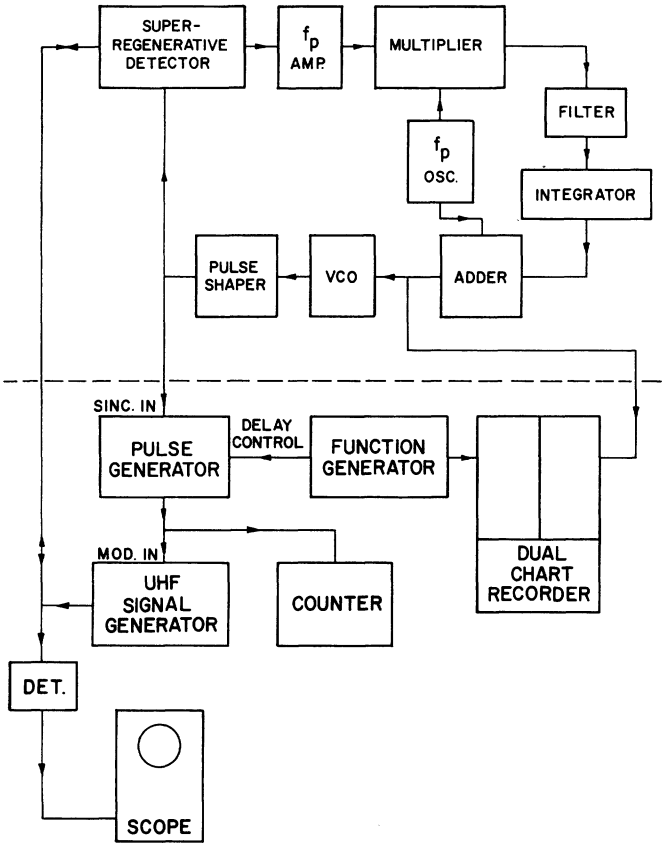

Fig. 12. Block diagram of the altimeter and the test system.

Two types of antennas were used in the flight test of the 403 MHz version. The first was a rhombic antenna with leg length of 3λ, and tilt angle of 58.5°. The measured gain of the rhombic was 9 dB over a half-wave dipole.

The second antenna was a two-element Yagi, with the parasitic element as the director, and a spacing of 0.13λ between the elements. The measured gain of the Yagi was 3.2 dB over a half-wave dipole.

The laboratory test system appears in Fig. 12. The quench pulses were used to trigger a delayed pulse generator. These delayed pulses modulated a UHF signal generator whose output simulated the returned pulses. The output of the signal generator was connected to the input (and also output) of the superregenerative detector. This point was also connected to a detector in front of an oscilloscope. This scope was used to monitor the shape of the envelope of the transmitted pulses. The counter was used to count the repetition rate. The chart recorder measured the input voltage to the VCO. These two pieces of equipment were used to measure steady-state error and noise error. The chart recorder was also used to measure the loop bandwidth. To do this the delay was modulated by a very low-frequency (VLF) sine wave from the function generator. The dual chart recorder recorded both the sinusoidal signal at the output of the function generator, and the resulting sinusoidal signal at the output of the integrator. By comparing these two amplitudes at several frequencies

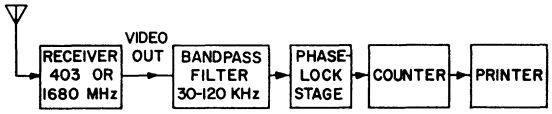

Fig. 13. Block diagram of the ground station.

it was possible to draw the frequency response of the loop.

The amplitude of the simulated returned pulse was controlled by the output attenuator of the signal generator. It should be noted that changes in the returned pulse amplitude affected mainly the signal-to-noise ratio but also the gain of the loop, as the gain includes in it the second derivative of the envelope of the returned pulse.

Without any sweeping voltage, the altimeter starts to lock when the peak of the returned pulses is as low as 2.5 μV. With proper sweeping voltage (the entire range is scanned in 25 seconds), good locking (with mean time to unlock on the order of hours) is achieved with 8 μV input. However, the steady-state error is about 0.3 percent and the noise error 0.1 percent. With an input of 15 μV the total error is below the specified 0.03 percent.

Data reception on the ground is accomplished by receiving and counting the pulse repetition rate. Two methods are used: either 1) receiving the altimeter RF pulses directly; or 2) modulating the VCO, or quench pulses, on the radiosonde signal and receiving those. In both cases the receiver at the ground station was to have IF and video bandwidth capabilities equal to the highest repetition rate, which, in our case is about 100 kHz. The video output of the ground station receiver is thus a carrier varying slowly between 50 to 100 kHz. From the altimeter loop bandwidth we can deduce that the information bandwidth is less than one hertz. This type of signal naturally requires a phase-lock demodulator for best reception. Yet, despite the fact that the information bandwidth is 1 Hz, the phase-lock loop bandwidth has to be at least f_p, which is the perturbation frequency. The block diagram of the ground station appears in Fig. 13.

Flight Tests

As of March 31, 1969, a total of 4 flight tests were conducted. The purpose of the first one was to test the principle of operation, i.e., whether the altimeter would lock to the returned pulse, and how it would follow changes in altitude. This test was conducted on June 5, 1968, over Lake Michigan, from an aircraft provided by the National Center for Atmospheric Research, Boulder, Colo. The antenna used in the flight was a slotted dipole. Excellent locking was achieved up to the maximum altitude of the flight—8000 feet. During the 30-minute period over the lake, locking was not lost even once, unless by external disturbance by the operator. Readings of altitude were performed by reading the period

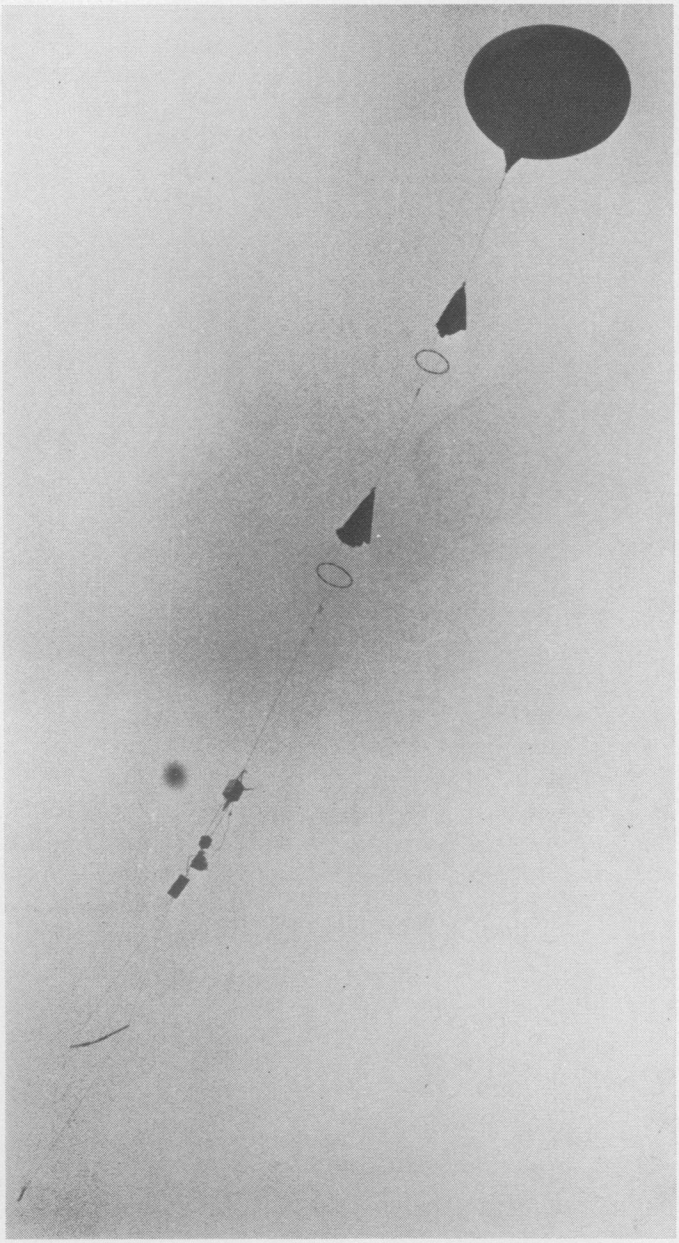

Fig. 14. The balloon payload.

of the quench pulses with a portable oscilloscope. These readings were compared to readings of the pressure altimeter of the aircraft. With these rather limited facilities, readings on the tested altimeter were within ± 0.1 percent of the readings on the aircraft altimeter.

The height of the waves in Lake Michigan at the time of the above test was 3 feet according to the U. S. Coast Guard. En route back to Madison, the altimeter was still locking over land at an altitude of 8000 feet. However, locking was lost from time to time.

In the second flight test, the radio altimeter, together with a 1680 MHz radiosonde, was carried aloft by a standard 1200 gram neophrene balloon. The altimeter was launched from Madison, Wis., on November 22, 1968, and landed on a farm in Kenosha County, Wis. The rhombic antenna design was used in this flight. Power was drawn partly from the radiosonde battery (100 V and 6 V) and partly from additional water activated batteries (+12 V and −12 V). The balloon and the payload of this flight are shown in Fig. 14 a few seconds after being launched. The items from the top are: the balloon, two parachutes, radiosonde, a package with the additional batteries, a bag with a letter to the finder, the radio altimeter, and the rhombic antenna. As far as altitude measurements were concerned, this flight was almost a complete failure. There was only a short intermittent period of about 10 minutes when locking was achieved. This happened at an altitude of 22 km which is rather high for an overland flight. The reason behind this is that, due to poor temperature and voltage compensation, the sweeping rate of the altimeter was decreasing, and at very slow sweeping rate its sensitivity improves considerably. Despite the failure in achieving meaningful altitude readings, many helpful conclusions were drawn from this test. The major one was that the method of counting the altimeter pulses directly from the video signal at the ground station receiver was inadequate as a result of noise. This conclusion led to the addition of a phase-locked stage between the receiver and the counter. Another observation from this test was that the radiosonde signal on 1680 MHz was stronger than the altimeter signal on 403 MHz. This was probably due to: 1) the fact that the altimeter antenna was pointing down while the radiosonde antenna was pointing sideways toward the ground station; 2) the higher frequency of the radiosonde which allows the use of a dish antenna at the ground station, with considerably more gain than the helix antenna used for the 403 MHz signal. This resulted in a change in which the altimeter VCO square wave is used to modulate the radiosonde 1680 MHz transmitter. Thus, the height information is transmitted to the ground through a better channel.

The third and fourth tests were also balloon flight tests. They were launched near Sturgeon Bay, Wis., on March 14 and 15, 1969. The balloon flight test of March 14, 1969, was the most successful thus far, and will be described in detail. In this test the altimeter utilized a 2-element Yagi antenna; the electronics and the antenna are shown in Fig. 15. The electronics weighs 130 grams and the antenna 40 grams. Separate batteries were used in this flight, so that the only electrical connection between the altimeter and the radiosonde is that of the altimeter VCO square wave, which modulated the radiosonde transmitter. Thanks to a calm day, we were able to use a very long string (about 300 feet) between the balloon and the payload, and thus minimize the pendulum motion. After burst of the balloon, the payload landed at Norwalk, Mich., and was returned to us. From that we can conclude that most of the flight, and certainly all of the ascent, took place over Lake Michigan.

Height readings started 7 minutes after launch at an altitude of 1.8 km, and were received for the next 57 minutes at intervals of 2 seconds, until the balloon

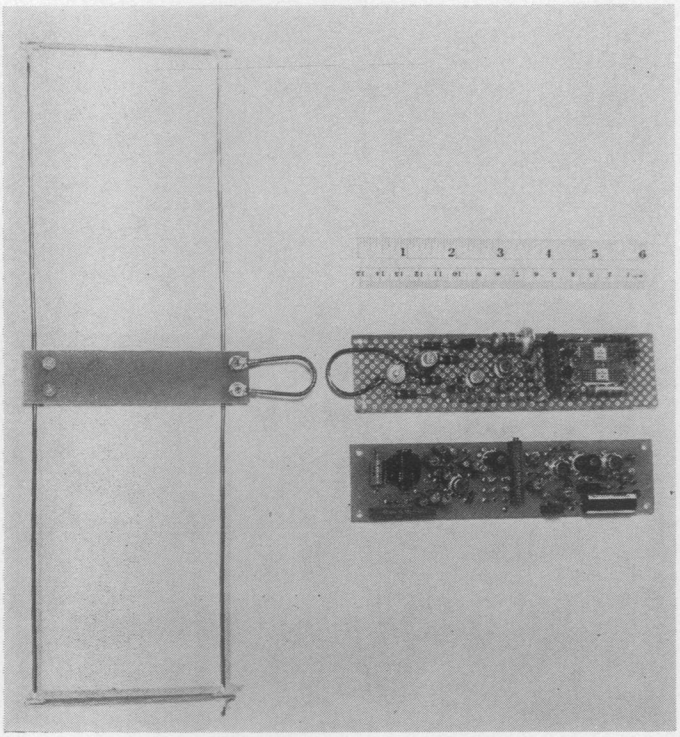

Fig. 15. The radio altimeter layout.

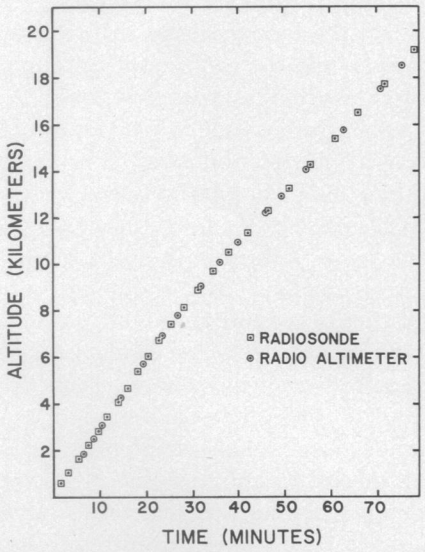

Fig. 16. Altimeter height readings, compared to radiosonde height calculations, during the ascent of March 14 1969, Sturgeon Bay, Wis.

reached an altitude of 16 km. In the next 10 minutes the altitude readings became intermittent. Yet, good data were still available up to an altitude of 18.5 km, where the signal-to-noise ratio of the radio altimeter decreased to the point that no locking could be achieved.

In Fig. 16 several height measurements of the altimeter in intervals of about 5 minutes, during the ascent, are compared to heights calculated from the temperature and pressure data received from the radiosonde. These calculations utilize the hydrostatic equation (1), where the layers Δh were between every fifth contact on

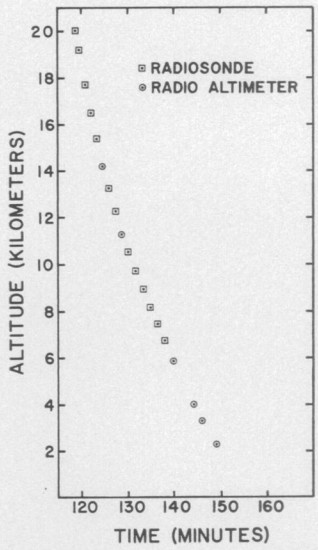

Fig. 17. Altimeter height readings, compared to radiosonde height calculations, during the descent of March 14, 1969, Sturgeon Bay, Wis.

the radiosonde baroswitch. Calibration of the baroswitch was performed two weeks before the flight.

Height readings were received also on the descent, starting at an altitude of 14.2 km and continuing down to 2.2 km, although the readings were intermittent during portions of the descent. Some of these readings are compared to the radiosonde data in Fig. 17.

The altimeter data, which is actually the repetition rate f_q, were recorded in a form of a 5-digit printer output every 2 seconds. The height h is related to the repetition rate through the following equation:

$$h(\text{meters}) = \frac{c(\text{m}/\mu\text{s})}{2}\left[\frac{n \times 10^6}{f_q(\text{Hz})} - \tau(\mu\text{s})\right] \qquad (21)$$

where τ is the width of the quench pulse (which was 1.40 μs in this case), c is the velocity of light in meters per μs, and n is the subharmonic number.

Altogether, more than 2500 meaningful height measurements were recorded during the flight. It is impossible to present all the data in this report. We chose the 200 mbar level, which happens to be the midrange of this flight, and presented readings received from the altimeter over a period of two minutes during the ascent (Fig. 18) and one minute during the descent (Fig. 19). From Fig. 18 we note that the 2 seconds rms value of the random error, at an altitude of nearly 10 km, is smaller than 7 meters. This is an error of 0.07 percent.

There is no simple way to measure the absolute measurement error of the radio altimeter. The altitude, as calculated from the radiosonde pressure and temperature readings, is certainly less accurate. However, it should be pointed out that the difference between the radiosonde calculated height and the altimeter measured height was smaller than 100 meters throughout the entire flight.

The major difference in the fourth test was that the

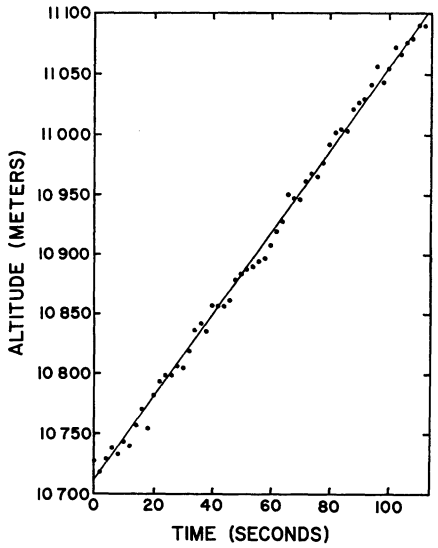

Fig. 18. Sample of the radio altimeter height readings over a two minute period during the ascent of March 14, 1969, Sturgeon Bay, Wis.

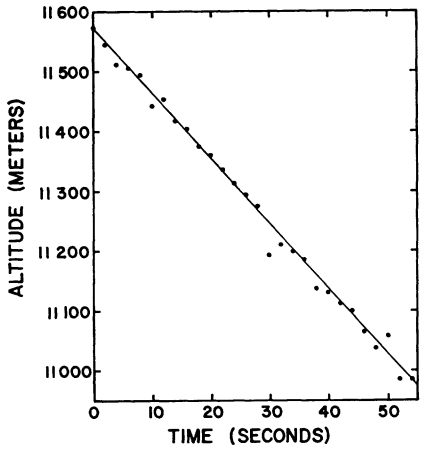

Fig. 19. Sample of the radio altimeter height readings over a one minute period during the descent of March 14, 1969, Sturgeon Bay, Wis.

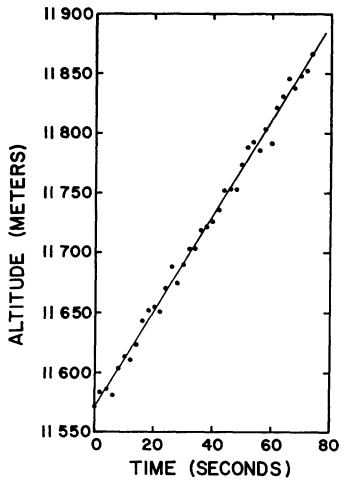

Fig. 20. Sample of the radio altimeter height readings over about a one minute period during the ascent of March 15, 1960, Sturgeon Bay, Wis.

rhombic antenna was used. This flight terminated near Williamsburg, Mich. From this we can conclude that nearly two-thirds of the flight took place over the lake. The rhombic antenna was much bulkier than the Yagi, and matching it to the altimeter indoors was more complicated. Immediately after the launch, we could see by monitoring the RF pulses that it was not matched properly. As a result, the RF pulse was too wide and the sweeping rate too fast. Both have the overall effect of reducing the altimeter sensitivity. Despite this, and probably due to the higher gain of the antenna, meaningful height measurements were received from the thirteenth minute until the 65th minute of the flight, when the altitude was 16.8 km. In Fig. 20 we give the readings received from this flight during a period of about one minute near the 200 mbar level. Contrary to the previous flight, the altimeter did not function at all during the descent. A possible explanation for this could be that the rhombic antenna became tangled during the descent. Both this and the problems involved in preflight matching of the rhombic lead to the conclusion that the Yagi is more suitable. An addition of several directors would give it the necessary gain to achieve the required range of 30 km.

Summary and Conclusions

This paper has described the basic theory and design of an accurate pulse radar altimeter. This altimeter is simple and light enough to be carried aloft by a regular sounding balloon.

Altitudes up to 20 km were measured in balloon flight tests over Lake Michigan. At midrange the rms value of the random error was smaller than 0.07 percent. Height readings were available every 2 seconds during both the ascent and the descent of the balloon payload. Small improvements in the antenna would easily increase the operational range to 30 km.

When flown together with a radiosonde, this altimeter can improve the accuracy of atmospheric sounding. It can also replace either temperature or pressure measurements through the use of the hydrostatic equation. This application is most attractive for high altitude radiosondes where the pressure resolution is inadequate, and for dropsondes when the thermistor time constant becomes too long compared to the rate of descent. Together with a pressure measurement, this altimeter can supply the reference pressure which is essential in the proposed system of floating superpressure balloons.

In addition to accurate height measurements, the altimeter can lock to an artificial return rather than its own returning pulse and, thus, measure distance to a transceiver. This can add wind measurement to its possible applications. All these applications should be studied and developed.

Further improvements in circuitry should also be considered. These may include: solid state RF stage at 1680 MHz, combining the several ICs into one package,

and elimination of all components that do not meet the design goals for air safety consideration.

Acknowledgment

I wish to extend sincere thanks to Prof. F. G. Stremler for his advice and encouragement during the work on this research. Very special thanks are due to Prof. V. E. Suomi, who brought me from Israel, suggested this research topic, gave me the facilities of his Space Science and Engineering Center, and was himself an endless source of ideas and inspiration. Thanks go also to C. Blair for his fine work on constructing the various models of the radio altimeter, and for his many suggestions concerning circuitry design. The help of J. Maynard in performing the flight tests was especially valuable.

References

[1] G. O. Hall, "Superregenerative receivers," in *Microwave Receivers*, M.I.T. Rad. Lab. Ser. New York: McGraw-Hill, 1948, ch. 20.
[2] V. W. Eveleigh, *Adaptive Control and Optimization Techniques*. New York: McGraw-Hill, 1967, ch. 9.
[3] A. R. Edison, R. K. Moore, and B. D. Warner, "Radar terrain return measured at near-vertical incidence," *IRE Trans. Antennas and Propagation*, vol. AP-8, pp. 246–254, May 1960.
[4] R. K. Moore and C. S. Williams, Jr., "Radar terrain return at near-vertical incidence," *Proc. IRE*, vol. 45, pp. 228–238, February 1957.
[5] M. Schwartz, "Effects of signal fluctuation on signal detectability," in *Modern Radar*, R. S. Berkowitz, Ed. New York: Wiley, 1965, pp. 130–140.
[6] A. J. Viterbi, *Principles of Coherent Communication*. New York: McGraw-Hill, 1966, chs. 1–4.
[7] F. M. Gardner, *Phaselock Techniques*. New York: Wiley, 1966.
[8] D. K. Barton, *Radar System Analysis*. Englewood Cliffs, N. J.: Prentice-Hall, 1964, pp. 25–29.
[9] N. Levanon, "Balloon-borne radio altimeter," Ph.D. dissertation (Appendix B), University of Wisconsin, Madison, April 1969.

Theory and Noise Dynamics of the Delay-Locked Loop

ALLEN G. LINDGREN, MEMBER, IEEE, ROBERT F. PINKOS, MEMBER, IEEE, AND MILTON E. SCHUMACHER, STUDENT MEMBER, IEEE

***Abstract*—The delay-locked loop, with its applications to radar, sonar, seismic propagation studies and interferometry, is a basic geoscience instrument. This paper presents the theory and investigates the performance of the delay-locked loop with and without clipping in a noisy (Gaussian) environment. The signal processing required to produce an appropriate error characteristic is derived and the significant factors affecting system operation are presented in the form of a normalized model. When signal clipping is introduced, a delay tracking system is shown to result that is self-adjusting and capable of providing near optimum performance in varying signal/noise environments. Application is made to the tracking of a research submersible randomly exploring the ocean floor. The dual problem of continuously tracking the peak spectral frequency of a random process exhibiting a slowly changing resonance is also considered.**

I. Introduction

THE ACCURATE determination of the delay difference between two versions of a common signal is necessary for precision angular tracking systems and distance measurement systems based on the propagation properties of the environment. This paper reviews the theory of the delay-locked loop (DLL) and presents a detailed analysis of the system's ability to achieve delay estimates in the presence of interfering noise. The delay-locked loop is a closed-loop system as shown in Fig. 1 where, in normal applications, the intentionally introduced, controllable delay estimate τ_b tracks the environmental delay τ_a. The role of the delay error generator (DEG) is to provide an indication of the delay difference $\tau_e = \tau_a - \tau_b$ between the two channels. The output of the DEG is smoothed by a loop filter which, in turn, commands changes in the controllable delay to reduce the delay difference to zero. When the system is properly operating, τ_b closely tracks or is "locked on" τ_a and provides an estimate of the desired delay.

Applications of the DLL include active ranging (radar and sonar), and angle tracking (interferometry). For the active ranging case $r_b(t)$ is a noise-free replica of $s(t)$; while in passive and active angle tracking applications, noise is present in both channels. A basic difference between the active and passive tracking problems, in addition to signal levels, is that the passive delay error generator must utilize the properties of the radiated target noise to produce a quantity dependent on the delay error τ_e. Unlike the active situation, where the signal $s(t)$ may be designed to optimize the shape of the error characteristic, the properties of the emitted target signal are normally not under the influence of the synthesist.

Manuscript received May 28, 1969; revised August 5, 1969. This work was supported in part by the U. S. Naval Underwater Weapons Research and Engineering Station, Newport, R. I., under Contract N00140-68-C-0447.

The authors are with the Department of Electrical Engineering, University of Rhode Island, Kingston, R. I. 02881.