VIII. FREQUENCY MODULATION

Prof. E. J. Baghdady
Prof. J. B. Wiesner
L. C. Bahiana
J. W. Conley
K. P. Luke
R. J. McLaughlin
R. E. Mortensen
G. J. Rubissow
F. I. Sheftman
R. H. Small
D. D. Weiner

A. ANALYSIS OF FM TRANSIENTS

A careful study of the conditions for low-distortion transmission of FM signals through linear filters suggests that, for purposes of quasi-stationary response, the most important characteristic of a frequency modulation is the maximum rate at which it tends to sweep the frequency of the carrier. Indeed, this maximum rate of change of instantaneous frequencies offers an excellent basis for classifying frequency modulations in the analysis of filter response. On this basis, three types of possible frequency changes may be recognized: (a) those that are relatively slow on the time scale of the system under consideration; (b) those that are so fast (on the time scale of the system) that they are essentially abrupt; (c) those that occur at a rate that is comparable to the speed with which the system response can build up and decay.

The analysis of filter response to an FM excitation can be carried out on an instantaneous-frequency basis, provided that

$$\epsilon = \frac{1}{2}\left|\frac{d\omega_i}{dt}\right|_{max} \cdot \left|\frac{Z''(j\omega_i)}{Z(j\omega_i)}\right| \ll 1 \qquad (1)$$

where $\omega_i(t)$ = instantaneous frequency of excitation, and $Z(j\omega)$ is the system function. The quantity ϵ has the significance of being an upper bound on the relative error in approximating the complex amplitude of the response by the quasi-stationary term only (Technical Report 332, in press, and Quarterly Progress Report of Jan. 15, 1958, p. 65). Equation 1 provides the criterion for deciding when $d\omega_i/dt$ is sufficiently slow on the time scale of the system to justify the use of the instantaneous-frequency method for analyzing the system response. In other words, only when this condition is satisfied can the system function be assumed to change instantaneously with $\omega_i(t)$.

When the frequency change of the excitation is essentially abrupt, we have another situation in which the analysis can often be reduced to a greatly simplified pictorial

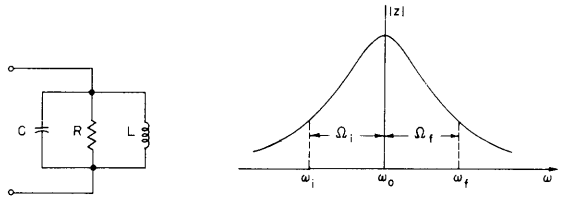


Fig. VIII-1. High-Q resonant circuit and definition of symbols.

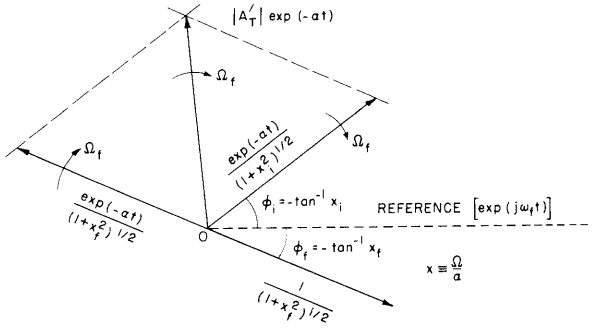


Fig. VIII-2. Phasor representation of response at t = 0+.

form. We shall illustrate this situation by considering the response of a single-tuned high-Q circuit to a frequency step.

With reference to Fig. VIII-1, assume that the excitation is described by

$$i(t) = \exp j\left\{\int^{t} [\omega_i + u_{-1}(t)(\omega_f - \omega_i)]\, dt + \phi_o\right\} \tag{2}$$

where $u_{-1}(t)$ is the unit-step function. The impedance of the high-Q single-tuned circuit can be written as

$$Z(j\Omega) = \frac{1}{1 + j\frac{\Omega}{a}} \qquad \Omega = \omega - \omega_o \tag{3}$$

where $a = 1/(2RC)$, and ω_o is the frequency of resonance. Before $t = 0$, the filter response is assumed to be in a steady-state condition. After $t = 0$, the response is made up of two sinusoids – one at center frequency and one at $\omega = \omega_f$. The sinusoid at the center frequency dies out with time, whereas the sinusoid at $\omega = \omega_f$ represents the steady-state component that corresponds to the new excitation at $\omega = \omega_f$. If e(t) is the filter response, then

$$e(t) = \begin{cases} Z(j\Omega_i) \exp(j\omega_i t) & \text{for } t \leq 0 \\ A_T' \exp(-at) \exp(j\omega_o t) + Z(j\Omega_f) \exp(j\omega_f t) & \text{for } t \geq 0 \end{cases} \tag{4}$$

Since the two expressions must be identical at $t = 0$, we have

$$A_T' = Z(j\Omega_i) - Z(j\Omega_f)$$

If we take $\exp(j\omega_f t)$ as a reference, then at $t = 0+$ the response is completely described by the phasor diagram shown in Fig. VIII-2.

Usually, the frequency deviations from the final frequency ω_f are of interest. Therefore, the analysis is best carried out by taking the final steady-state response as reference and normalizing with respect to $Z(j\Omega_f)\exp(j\omega_f t)$. The result of this operation is the simplified representation

$$e'(t) = \frac{e(t)}{Z(j\Omega_f)}\exp(-j\omega_f t)$$

$$= 1 + A_T\exp(-\alpha t)\cdot\exp[j(\phi_T - \Omega_f t)] \qquad (5)$$

which is illustrated by the phasor diagram of Fig. VIII-3, where the symbols are also defined.

The phasor model of Fig. VIII-3 can be used to explain all of the amplitude and frequency transient phenomena that are observed in the laboratory (see Section VIII-B). Moreover, all of the significant information (for example, the magnitudes and times of occurrence of overshoots and undershoots, the conditions that give rise to amplitude nulls, and rise times) can be quickly obtained from this diagram by inspection. Indeed, the simplification is so great that the problem becomes almost trivial!

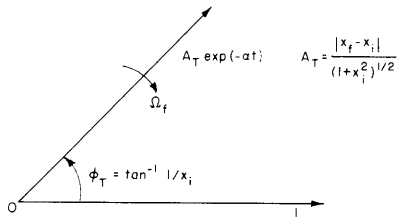


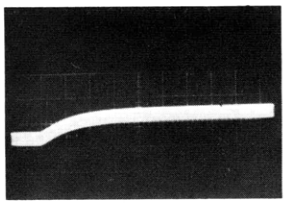
Fig. VIII-3. Simplified phasor diagram of filter response.

With a more complicated filter, any initial energy that might be stored in the system before $t = 0$ will be dissipated at the natural frequencies (normal modes) of the system after $t = 0$. At $t = 0+$, the term that represents the final steady-state response will be cancelled out by a term whose components die out with time at the natural frequencies of the system. Once the component solutions have been determined, the determination of the instantaneous-amplitude and instantaneous-frequency behavior of the total response is best carried out by solving a multisignal interference problem. In the case of band-pass filters whose pole patterns display symmetry about some frequency, significant additional simplifications are possible.

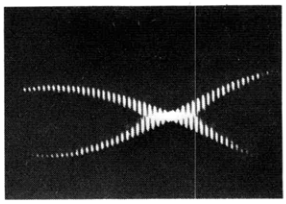
E. J. Baghdady

B. EXPERIMENTAL STUDY OF FM TRANSIENTS

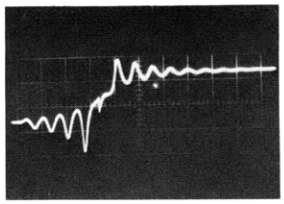
Oscillograms of the instantaneous frequency of the response of the single-tuned circuit to a series of frequency steps that start at the mid-band frequency of the filter are shown in Fig. VIII-4. For x_f less than 0.80, the response resembles a rising exponential whose rise time is approximately independent of x_f (see Fig. VIII-4(a)).



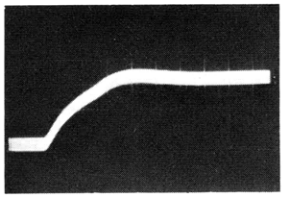
(a). $x_i = 0$; $x_f = 0.40$



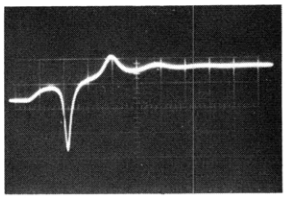
(f). $x_i = 0$; $x_f = 3.92$



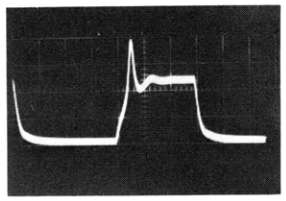
(k). $x_i = 0$; $x_f = 11.80$



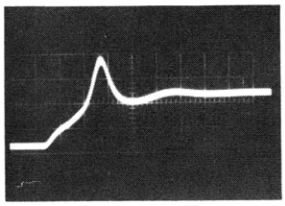
(b). $x_i = 0$; $x_f = 1.20$



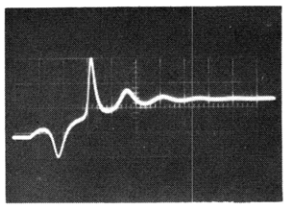
(g). $x_i = 0$; $x_f = 4.00$



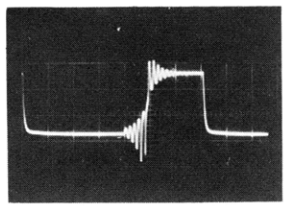
(l). $x_1 = 0$; $x_2 = 2.40$



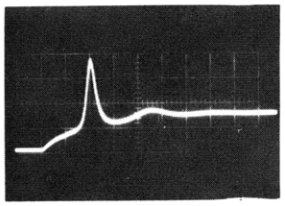
(c). $x_i = 0$; $x_f = 2.40$



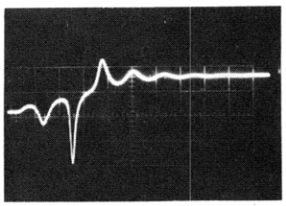
(h). $x_i = 0$; $x_f = 5.40$



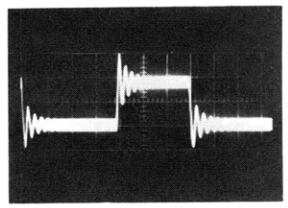
(m). $x_1 = 0$; $x_2 = 14.08$



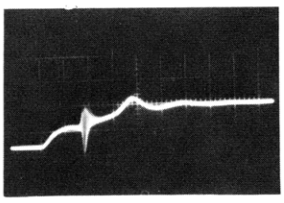
(d). $x_i = 0$; $x_f = 3.00$



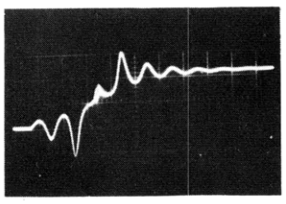
(i). $x_i = 0$; $x_f = 6.20$



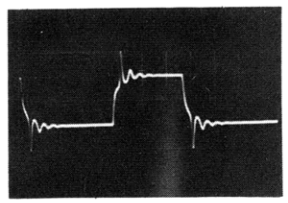
(n). $x_1 = 7.04$; $x_2 = 7.84$



(e). $x_i = 0$; $x_f = 3.92$



(j). $x_i = 0$; $x_f = 7.80$



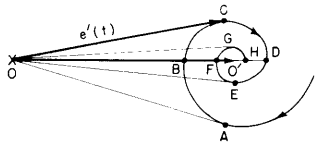
(o). $x_1 = -5.24$; $x_2 = +5.24$

Fig. VIII-4. Oscillograms of instantaneous frequency of response of the single-tuned circuit to a series of frequency steps that start at mid-band frequency of the filter.

A small overshoot is first noticeable for $x_f = 0.80$. As x_f becomes greater than 0.80, the magnitude of the overshoot increases with a corresponding decrease in the rise time of the response and the time of maximum overshoot. An inflection point is also apparent on the leading edge of the transient (see Fig. VIII-4(b), (c), (d)). For $x_f = 3.92$, the overshoot is caught in the act of turning into an undershoot, and, at the instant of transition, the instantaneous amplitude of the response drops to zero. These two effects are shown in Fig. VIII-4(e) and (f), respectively. For $x_f = 4.00$, the "overshoot" has definitely become an undershoot (see (g) of the figure). With a further increase in x_f, the magnitude of the first undershoot decreases in value, whereas the magnitude of the second overshoot increases (see (g) and (h)). For $x_f = 6.00$, the second overshoot becomes an undershoot (see (i)), and the instantaneous amplitude of the response again drops to zero. For larger values of x_f, the magnitude of the second undershoot decreases in value. The third, fourth, and successive overshoots behave exactly the same as the first and second. Note that the undershoots, when they occur, always precede the overshoots and that, for a given x_f, successive undershoots are always larger than the preceding ones, whereas successive overshoots are always smaller. Note also that the frequency of ringing in all cases approximately equals Ω_f (see (j) and (k)).

In Figs. VIII-4(l) and (m), the leading edge of the square wave corresponds to a frequency deviation away from the center frequency ($x_i = 0$) of the tuned circuit; the trailing edge corresponds to a deviation towards the center frequency ($x_f = 0$). When $x_f = 0$, the response is always exponential and has no overshoot, regardless of the value of x_i. Figure VIII-4(n) shows a response in which the overshoots have no inflection point in the leading and trailing edges. Figure VIII-4(o) indicates that when $x_i = x_f$ the leading and trailing edges of the square-wave frequency response have identical rise times, percentage overshoots, and times of maximum overshoot.

The vector model of Fig. VIII-3 of Section VIII-A can be used to explain all the phenomena observed in the FM transient oscillograms. Consider, for example, the



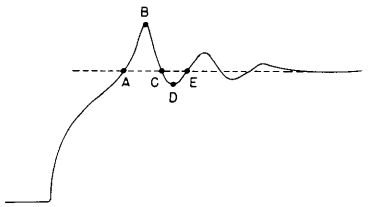


Fig. VIII-5. Behavior of e'(t) that results in occurrence of instantaneous-frequency and instantaneous-amplitude transients in the tuned-circuit response. Note that O-O' is the unit vector.

situation shown in Fig. VIII-5. As the vector $A_T \exp(-\alpha t)$ decays exponentially, the vector e'(t) wobbles back and forth with a period approximately equal to $1/(2\pi\Omega_f)$. At point A, e'(t) is at rest, and the instantaneous angular frequency of e(t) equals ω_f. As e'(t) rotates towards B, the time rate of change of the phase angle, θ, increases and an overshoot in the instantaneous-frequency response of e(t) results. The maximum overshoot occurs at point B. The time rate of change of θ decreases as e'(t) rotates from B. At point C, e'(t) is again at rest. As e'(t) rotates towards D, $d\theta/dt$ continues to decrease and the instantaneous angular frequency of e(t) falls below ω_f. The angular frequency reaches its minimum value at point D. As e'(t) rotates from D, $d\theta/dt$ begins to increase once again. At point E, $d\theta/dt = 0$, and one cycle of ringing in the instantaneous-frequency response of e(t) is completed. The cycle is repeated for points E, F, G, and H and continues to repeat until $A_T \exp(-\alpha t)$ is negligible with respect to the unit vector. Since such points as B and F occur at increasingly greater distances from O, overshoots in the frequency response of e(t) occur with successively smaller magnitudes.

The occurrence of undershoots in the instantaneous frequency of the response is explained with the use of Fig. VIII-6. For the case shown, the vector $A_T \exp(-\alpha t)$ is sufficiently large with respect to the unit vector at $t = 0$, and the frequency Ω_f is sufficiently high to enable the tip of the e'(t) vector to encircle the origin, O, twice as $A_T \exp(-\alpha t)$ revolves at the angular frequency, Ω_f. Since the instantaneous frequency of the resultant of two signals of different frequencies always overshoots in the direction of the frequency of the stronger signal, the instantaneous frequency of e(t) undershoots at such points as A and B, which occur to the left of O, and overshoots at such points

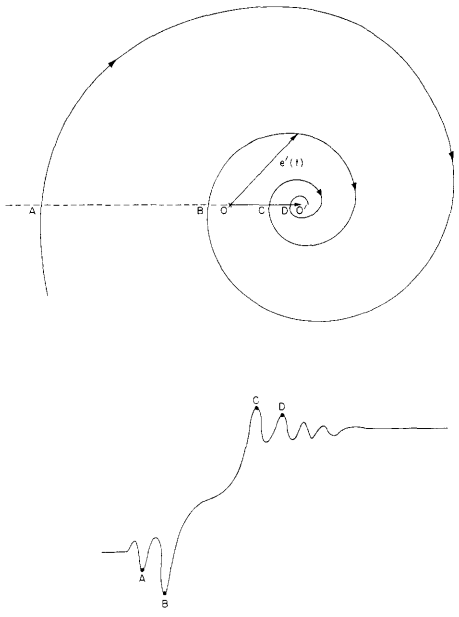


Fig. VIII-6. Behavior of e'(t) that results in occurrence of both undershoots and overshoots in the instantaneous-frequency response of e(t). Note that O-O' is the unit vector.

as C and D, which occur to the right of O. Note that the undershoots, if they occur, must always precede the overshoots; also, successive undershoots have magnitudes that increase in value, whereas successive overshoots have magnitudes that decrease in value. When the tip of the e'(t) vector passes through the origin, the magnitude of e'(t) equals zero, and the instantaneous frequency of the response is transitional between an overshoot and an undershoot.

When $\omega_f = \omega_o$, $\Omega_f = 0$ and the vector $A_T \exp(-\alpha t)$ is stationary with respect to the unit vector. Thus, the response to a frequency jump that terminates at ω_o consists of a smooth exponential rise caused by the exponential decay of $A_T \exp(-\alpha t)$. (See Fig. VIII-3 of Section VIII-A.) When ω_f approximately equals ω_o, Ω_f is very small. Thus, $\phi = \phi_T - \Omega_f t$ is approximately constant as $A_T \exp(-\alpha t)$ decays, and the response is again an exponential rise.

The inflection point that occurs, for certain values of Ω_f, in the leading edge of the frequency response is caused by the fact that the transients in the frequency-step response of the single-tuned circuit are the result of the exponential decay of $A_T \exp(-\alpha t)$, and the angular rotation of $A_T \exp(-\alpha t)$. When first one effect predominates and then the other, an inflection point is seen on the leading edge of the response. If Ω_f is small, only the first effect predominates, and no inflection point is observed. If Ω_f is very large, only the second effect predominates, and again, no inflection point occurs.

D. D. Weiner

C. DYNAMIC TRAP FOR THE CAPTURE OF THE WEAKER SIGNAL

An FM receiver that incorporates a dynamic trap system for capturing the weaker of two cochannel signals has been built and tested (1, 2). A summary of the results follows.

The capture performance of an FM receiver is brought out clearly by a test procedure that simulates the cochannel interference by the superposition of two carriers that have easily identifiable modulations. A suitable choice of modulation for facilitating capture measurements is a sinusoidal message of known frequency. The receiver performance can then be presented in terms of a plot of the fundamental component of stronger (or weaker) signal modulation frequency as a function of the weaker-to-stronger signal amplitude ratio, a; $a = E_W/E_S$. Generally a more complete presentation involves both weaker-signal and stronger-signal capture characteristics plotted on the same coordinates.

The capture characteristics of the dynamic-trap receiver are determined largely by two parameters that describe the attenuation characteristic of the trap. The more important parameter is the maximum trap attenuation factor, δ, which is defined as the maximum trap attenuation within the passband, normalized with respect to the

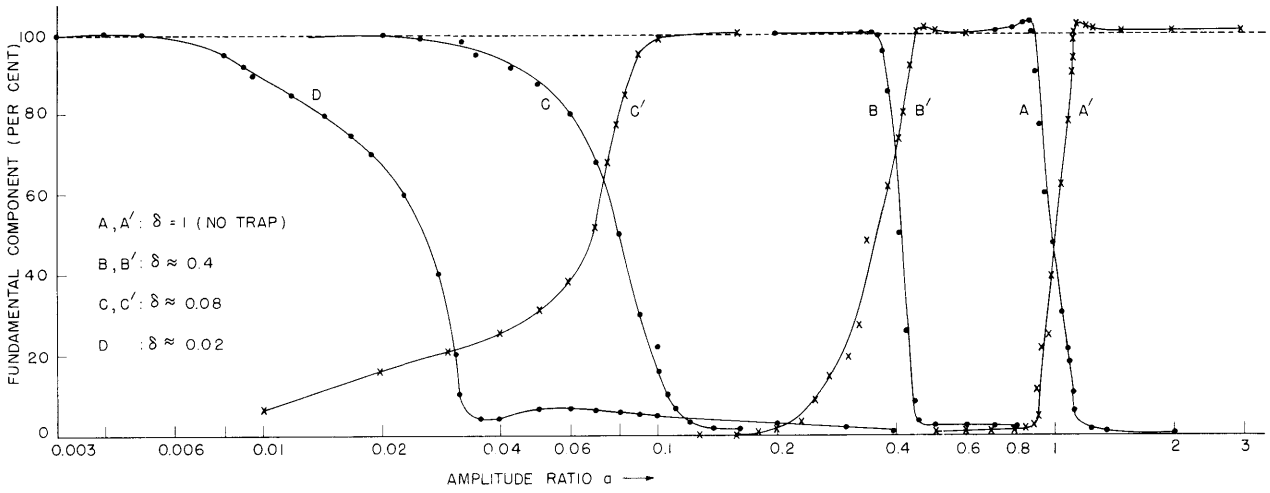


Fig. VIII-7. Capture characteristics for four values of δ; bandwidth is fixed at 15 kc. Curves A, A', B, B', C, C': E_S is modulated ±30 kc at 400 cps; E_W is modulated ±30 kc at 1 kc. Both signals are centered in the passband. Curve D: E_S is modulated ±35 kc at 400 cps and centered in passband; E_W is unmodulated and at edge of passband. (Primes denote weaker-signal capture characteristics.)

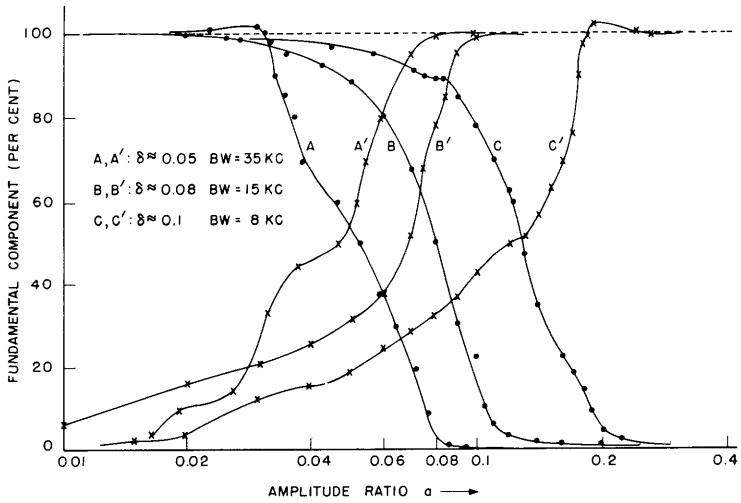


Fig. VIII-8. Capture characteristics for three values of bandwidth. E_S is modulated ±30 kc at 400 cps. E_W is modulated ±30 kc at 1 kc. Both signals centered in passband.

maximum amplitude response in the passband. The less important parameter is the trap bandwidth, BW. This corresponds to the difference between the half-power frequencies of the "inverted" frequency response of the trap. At those frequencies the relative passband gain is reduced to $[1 - (1 - \delta)/\sqrt{2}]$.

Figures VIII-7 and 8 show capture characteristics of the experimental system. In the measurements that led to curves A, B, and C of Fig. VIII-7 and curves A, B, and C of Fig. VIII-8, the two signals are at the center of the band, and each is modulated with a deviation of ±30 kc. The modulation frequencies are 1 kc for the weaker signal and 400 cps for the stronger signal.

Curves A and A' of Fig. VIII-7 show the observed capture characteristic of the receiver without the trap. The curves for the weaker and stronger signals are symmetric about the $a = 1$ line. The capture transition region is relatively narrow and is centered at $a = 1$.

Curves B, B', C, C', and D of Fig. VIII-7 show characteristics for the receiver with the dynamic trap for three values of δ and a fixed trap bandwidth of 15 kc. The best observed performance of the system is shown by curve D. Here, the weaker signal is unmodulated and sits at the edge of the passband. The stronger signal is centered in the middle of the band and is modulated through ±35 kc at 400 cps. The fundamental component of the stronger signal falls to a relatively small value when $\underline{a}$ is greater than 0.04. From this it is inferred that the output consists mostly of the weaker signal when $\underline{a}$ is greater than 0.04. The curves of Fig. VIII-7 clearly show that the capture transition region always occurs at $a \approx \delta$, and hence that the weaker signal is captured for values of $\underline{a}$ slightly greater than δ.

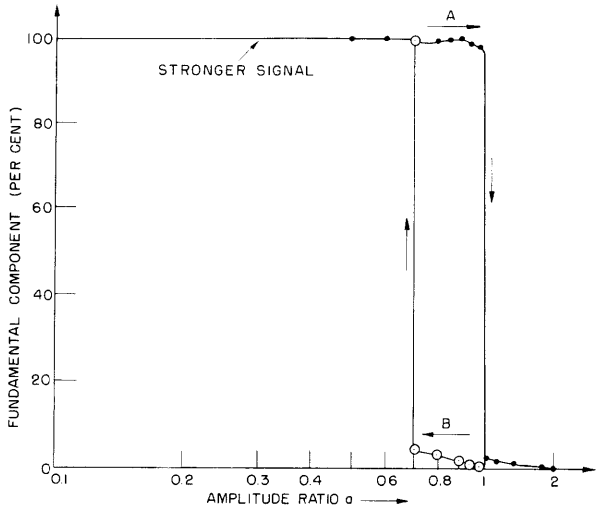


Fig. VIII-9. Preliminary stronger-signal capture characteristic for receiver with dynamic signal selector.

Figure VIII-8 shows capture characteristics for three different trap bandwidths. The value of δ was adjusted to the minimum attainable for a given bandwidth. It ranged from approximately 0.05 to 0.1. We see that the bandwidth primarily affects the detailed shape of the capture characteristics. The most abrupt capture transition region occurs for an intermediate bandwidth (3).

These results indicate the feasibility of a dynamic trap system for capturing the weaker signal. For best performance and stability, the trap should be designed for a minimum value of δ and a large bandwidth.

Some experimental work also has been done to test a dynamic selector (1) that improves the capture of the stronger signal. The demodulated output from the receiver is fed back to deviate a peaked filter in the i-f stages, so that its center frequency follows the instantaneous frequency of the stronger signal. A preliminary capture characteristic is shown in Fig. VIII-9. Both signals were centered in the passband. The weaker signal was unmodulated; the stronger was modulated through ±30 kc at 400 cps. The peaked filter had a Q of 88. As $\underline{a}$ is increased from 0 (curve A), the stronger signal is captured 100 per cent when $\underline{a}$ is less than 1.02. As $\underline{a}$ is decreased from 2 (curve B), the stronger signal is recaptured only when $\underline{a}$ is less than 0.7. Thus, the system exhibits a hysteresis-like charcteristic. This system also shows promise of reducing impulsive interference.

G. J. Rubissow

References

1. E. J. Baghdady, Capture of the weaker signal, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., Oct. 15, 1957, p. 53.
2. G. J. Rubissow, Dynamic trap for capture of the weaker signal in FM interference, S.M. Thesis, Department of Electrical Engineering, M.I.T., January 1958.
3. Ibid., pp. 111, 117.

D. TRANSISTOR PULSE-CIRCUIT ANALYSIS

In the usual analysis of transistor pulse circuits, the transistor is replaced by a piecewise-linear equivalent circuit. The transistor operates as a charge-controlled device. An oversimplified model that takes the charge-controlled operation into account makes use of capacitors. From this approach the effects of transistor geometry and of the constituent semiconductor material parameters are not obvious, and the analysis often leads to results that lack physical significance.

In this investigation, the transistor is studied in a typical pulse-circuit application, a blocking oscillator. In the earlier phases of the investigation, experiments with junction diodes and driven-transistor circuits were necessary for deducing the fundamental

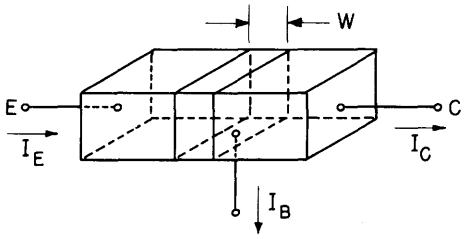


Fig. VIII-10. Transistor geometry.

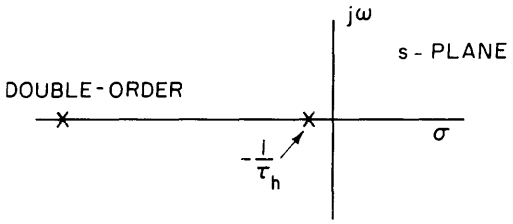


Fig. VIII-11. Approximate pole pattern for β(s) in the s-plane.

properties of the transistor as a circuit element. Theoretical analysis that assumes ideal geometry and material behavior is another way of deducing circuit performance. We used both an empirical and a theoretical approach and applied the results to the blocking oscillator. Although the study is primarily of academic interest, the results provide a basis for circuit design and an understanding of the important effects contributing to the performance of the device.

The theoretical part of the investigation consists of the solution of the time dependence of driving-point and current-transfer relations. With parallel-plane geometry (Fig. VIII-10), the assumption of linear diffusion and recombination of excess charge density of a minority carrier, and suitable approximations to the boundary conditions, the common-emitter transfer function for arbitrary currents in normal bias is given by

$$\beta(s) = \frac{I_c(s)}{I_b(s)} = \frac{1}{\cosh \frac{W}{L_h} (s\tau_h+1)^{1/2} - 1} \tag{1}$$

The s-plane representation of this expression is shown in Fig. VIII-11. The quantity τ_h is the effective recombination lifetime of bulk and surface properties of the semiconductor material. The transfer function β(s) can be simplified, without loss of generality except for very small portions of the transient, by taking the term corresponding to the pole closest to the origin, so that we have

$$\beta(s) = \frac{2\left(\frac{L_h}{W}\right)^2}{s\tau_h + 1} = \frac{\beta_o}{s\tau_h + 1} \tag{2}$$

This is a familiar form for the current-transfer ratio. Here β_o is the static current gain, and τ_h is the exponential time dependence. If a time constant τ_r is assigned to the transistor, experimental measurement of the recombination lifetime is unnecessary.

Other important effects are caused by the collector depletion layer, base-width modulation, and nonlinearity in the base driving-point characteristic. Fortunately, these are not important effects in many pulse circuits. Assumptions of an ideal current source representation at the collector and a purely resistive base driving-point impedance are completely justified in a large number of pulse circuits.

Switching from saturation, through normal bias, to current cutoff, and vice versa, is of interest in pulse circuitry. The state of the device is determined by the excess charge density of the minority carrier, which is a function of the applied current. Often an initial distribution of charge density (as found in saturation, for example) is known, and the applied base current can be found from circuit relations. For this situation, a useful relation is

$$Q(s) = \frac{Q_o + I_b(s)}{s + \frac{1}{\tau_h}} \tag{3}$$

where Q_o is the initial, stored, excess charge density in the base, $I_b(s)$ is the applied base current, $Q(s)$ is the time dependence of total stored excess charge in the base, and τ_h is the effective recombination lifetime.

If it is possible to assume that zero stored charge density is the condition for current cutoff, then it is possible to determine the saturation time. This assumption is justified in most pulse-circuit applications if the base drive is not greater than the forward current in the emitter.

From these theoretical considerations, a solution for the blocking oscillator shown in Fig. VIII-12 can be found. The collector voltage waveform of interest is shown in Fig. VIII-13. The important time relations, obtained directly from active circuit theory, are

$$T_1 = T_3 = \frac{\tau_r}{\left[\frac{n\beta - \frac{n^2 r_s \tau_r}{R_\ell RC}}{1 + \frac{n^2 r_s}{R_\ell}} - 1\right]} \tag{4}$$

$$\left[\frac{\beta}{n}\,\frac{1}{\frac{\tau_r}{RC}-1}\,\exp(-T_1/\tau_r)\right]\exp(-T_2/\tau_r) = \frac{r_b}{R_\ell} + \frac{T_2 r_b}{L_m} \tag{5}$$

In the derivation of T_1 the concept of a right half-plane pole is made. Correspondingly, a transient of incrementally small initial value is generated. The time relation is then obtained directly. In computing T_2, a maximum allowed collector current supported by

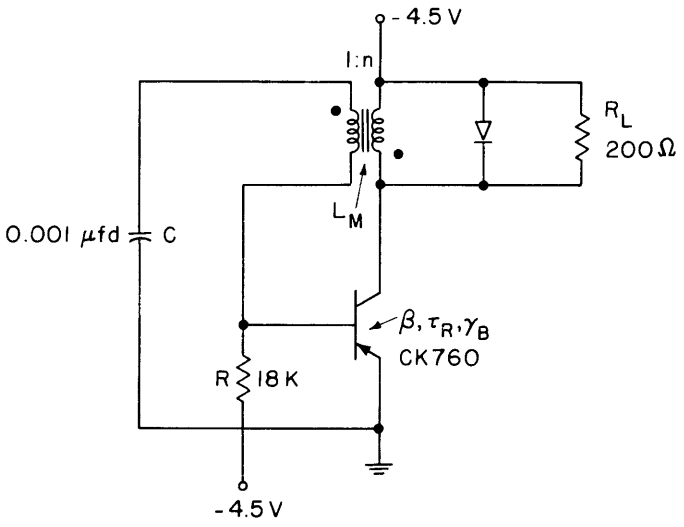


Fig. VIII-12. Circuit of experimental blocking oscillator. L_m = 840 microhenrys; r_e = 105 ohms; β = 20; τ_r = 0.86 μsec; n = 1.

a charge in the base is equated to the load current determined by the circuit. The transcendental nature of the expression for the pulsewidth is the result of the simultaneous contribution of increasing magnetizing current and recombination of excess charge density of the minority carrier in the base. Note that the storage effects contribute only to the width of the pulse. In the absence of this effect, a static current gain would have

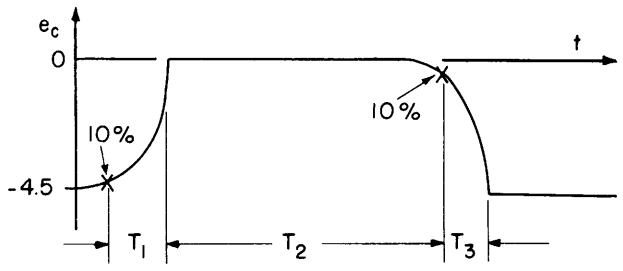


Fig. VIII-13. Collector voltage waveform.

to exist, and its magnitude would be

$$\beta \geq \frac{R}{R_\ell} = \frac{18,000}{200} = 90 \qquad (6)$$

The agreement of theory and experiment is good. Although the derived relations are not of general applicability, the method by which they are derived can be extended to circuits of increasing complexity, with a corresponding increase in mathematical difficulty.

J. W. Conley