JPL Publication 98-**

Communications and Control for Electric Power Systems:

Final Report

H. Kirkham

April 1998

Prepared for

Office of Energy Management Systems
United States Department of Energy

Through an agreement with

National Aeronautics and
Space Administration

by

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL Publication 98-**

Communications and Control for Electric Power Systems:

Final Report

H. Kirkham

April 1998

Prepared for

Office of Energy Management Systems
United States Department of Energy

Through an agreement with

National Aeronautics and
Space Administration

by

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible electronic image products. Images are produced from the best available original document.

Prepared by the Jet Propulsion Laboratory, California Institute of Technology, for the U.S. Department of Energy through an agreement with the National Aeronautics and Space Administration.

This publication reports on work performed under NASA Task RE-152, Amendment 203 and sponsored through DOE/NASA Interagency Agreement No. DE-AI01-79ET 29372 (Mod. A009).

OVERVIEW

THIS REPORT is a summary of some of the work done on the Communications and Control project, with particular emphasis on the achievements during the years 1986–1996. During those years, the project moved away from concern with dipersed storage and generation and its impact on power system operation (the team was responsible for studies in this area, and for making a power system simulator that included DSG), and became involved in more concrete work aimed at applying "high-tech" solutions to problems of power system communications and control.

This report covers work done at JPL on the following topics:

- the measurement of electric and magnetic fields, both ac and dc
- the use of optical power to supply low-power electronics
- the design of a fault-tolerant communication system designed for distribution automation
- a digital phase locked loop that allows the use of low-power transmitting electronics to recreate a good-quality signal at the receiver

In a report of this kind, only the results and highlights of the work are described. The reader who wants to see the entire development is urged to consult the many references to JPL reports and technical papers. The same need to consult other references exists for other work done on the project, including:

- a high voltage sensor based on the use of an electro-optic polymer
- the application of an artificial neural network (Kohonen's self-organizing feature map) to power system security and to power system stability.
- the integration of new control technologies for power generation and delivery
- the question of calibrating and interfacing low-energy instruments

all of which were the subjects of JPL reports or IEEE papers (the work on neural networks led to the project's main contributor on this topic producing a good deal of a new IEEE PES tutorial on the application of neural networks to power systems).

What follows is a mixture of old work and new words. Text from earlier reports has been re-edited and cut or augmented as appropriate for a final report. Because the means of production is new, most of the graphics are entirely new.

FOREWORD

POWER SYSTEMS might now include significant storage and renewable-resource generation if America had kept to its 1970's goal of energy independence. Such generators would produce unpredictable amounts of power, depending on the sun and the wind, in parts of the power system not accustomed to generation.

Operation of a power system with non-schedulable generators was known to be challenging. In anticipation of this, in the late 1970s the US Department of Energy began a project at the Jet Propulsion Laboratory aimed at easing the integration of what was then called dispersed storage and generation into the utility power system.

After I became its manager in 1984, the JPL project moved out of the business of dispersed storage and generation, and began to investigate the application of some advanced technologies to power system control and communication.

The project was not the first to use optically powered electronics, but it was the first to demonstrate its application to electric power. The technology of hybrid integrated circuits led to our development of a range of meters for the electric and magnetic fields of power lines. We were the first to measure the electric field of a dc line away from the ground. In response to a growing interest in distribution automation, the project developed a new set of network protocols aimed specifically at the communication needs of the low-voltage network.

These, and other developments, are described in this Final Report. It is sincerely hoped that the new competitive pressures on the industry will allow renewed, serious consideration of some of them.

PASADENA, CALIFORNIA
MAY 1994

HAROLD KIRKHAM

Section 1. ELECTRIC AND MAGNETIC FIELDS

SUMMARY

This section of the Final Report describes a fiber-based electrometer, two electronic space-potential electric field meters, and an isolated magnetic field meter.

One of the E-field meters was designed to measure dc fields, the others (an electrometer and a spherical-probe electronic system) ac fields. Both electronic systems use fiber optics to couple a small measuring probe to a remote readout device, so as to minimize field perturbation due to the presence of the probe. By using coherent detection, it has been possible to produce instruments whose operating range extends from about 10 V/m up to about 2.5 kV/cm, without the need for range switching on the probe. The electrical and mechanical design of both meters is described in detail. In dc fields, it was found that even the best fiberglass could not insulate the probe for long, and a solution using a small radioactive source was found.

Data from tests done in collaboration with the Electric Power Research Institute's High Voltage Transmission Research Facility (HVTRF)are described. An unexpected finding here was that the field enhancement caused by a person under a transmission line was much less if the person was wearing a hard hat.

The report also describes work done to measure the field near a high voltage dc bushing in a Swedish laboratory. The results obtained with the field meter indicate that a partially wetted insulator has uneven distribution of electric stress, a kind of behavior that has no counterpart in ac systems. It is no exaggeration to say that this kind of field enhancement would not have been widely expected.

Taken together, the results of measurements with the JPL meter—the pole leakage question, the reduced field over the individual with a hard hat and the field enhancemnt due to partial wetting—have lead to an improved understanding of the behavior of insulation systems stressed with direct voltage. One might generalize to say that dc insulation systems are controlled by surface effects, whereas ac insulation is controlled by capacitive voltage distribution. The implication is that dc insulation design and testing may have to be re-thought. Bucket trucks, for example, may be easier to test with dc, but the results would have no bearing on the

performance with ac energization.

A magnetic field meter that uses fiber optics to minimize electric field perturbation due to the presence of the probe and to provide electrical isolation for the probe is described. Such a meter could be used in a high field or high voltage environment. Power to operate the sensor electronics is transferred via an optical fiber, and converted to electrical form by a small photodiode array. The fundamental, the second and third harmonics of the field are filtered and separately measured, as well as the broadband rms level of the field. The design of the meter is described in detail and data from laboratory tests are presented.

1.1 ELECTRIC FIELDS

1.1.1 INTRODUCTION

For some while there has been interest in the effects of the electric fields associated with electric power transmission. As transmission lines were designed and built to operate at higher and higher voltages, their presence became more noticeable. A casual observer can hear the audible effect of rain on a high voltage alternating current transmission line, and sometimes the electric field itself can be sensed. A number of laboratories throughout the world have undertaken to investigate the effects, if any, of the fields associated with the transmission of electric power. These effects include the production of audible noise already alluded to, the generation of radio interference, and effects on biological systems.

These studies were conducted, in the main, to investigate the effects of alternating electric fields. Almost all the transmission lines in the world carry alternating current (ac), so this was a reasonable place to begin. However, under some circumstances, the transmission of power by direct current (dc) can be more efficient and more economical, and several large dc lines have been built. A natural extension of the work on ac field effects was to perform similar studies for dc fields.

It would be fair to stress, at this point, that the biological studies were faced with a very difficult problem. Early attempts to investigate the impacts of the electric fields associated with the transmission of electric power faced the difficulty that there were no models to help understand what might be going on. No satisfactory mechanism for biological impact had been proposed. No clear-cut experimental result had demonstrated a biological effect. The work was definitely of a pioneering nature.

Biological experiments are among the most difficult ones to perform because many variables, aside from the ones that are the subject of the study, can affect the outcome of the experiment. Under such circumstances repeated studies, with controls and carefully monitored conditions, are the only way that an increased understanding can be gained. The method is simple in principle: experimental subjects are exposed to electric fields for some amount of time, and later examined for possible effects. The subjects might be linemen working near high voltage equipment, or farm animals grazing under real transmission lines, or small animals in indoor test apparatus.

Naturally, since electric field is the parameter under investigation, it is essential to be able to measure or calculate the electric field that the subject experiences. The branch of electrical engineering known as field theory can describe the electric field at any point in space provided the correct "boundary conditions" are given. Closed form mathematical solutions to field

problems are given in all the standard texts for boundaries that have an appealing degree of symmetry. In the case of the animals that graze under transmission lines, it is not possible to transform the problem to one of suitable symmetry. Closed form solutions are not possible, and even numerical solutions are difficult. The presence of space charge considerably complicates the situation.

A much more satisfactory solution is the experimental measurement of the electric field environment as part of the study. In this way, the effects of variables that are not well understood (atmospheric humidity, temperature, etc.) can be accommodated along with those parameters that are known to be important (such as voltage and conductor-ground spacing). Appropriate instrumentation is essential.

When this work began, a number of instruments had been developed that could measure the value of the electric field under an ac transmission line, either at the surface of the ground or above it. Instruments with sensing elements that are isolated are known as free-body or space-potential instruments, because they permit the measurement of the field at the potential of the space in which they are used, without affecting that potential in any significant way. Electric field sensors typically have conductive electrodes and some form of cabling connecting them to the measurement electronics. This arrangement can cause considerable distortion of the unknown field if the measurement is to be made above the ground plane. In one implementation for ac measurements (Deno and Zaffanella, 1975) the sensor is self-contained, and furnished with a scale large enough that an operator a meter or two away can read it visually. The device is calibrated in a uniform field in a large test chamber. However, the probe is relatively large, and it distorts the field nearby, introducing an error if the field is non-uniform (as it would be near another conductor.) There is clearly some advantage to a small, all-dielectric sensor.

In other approaches, Feser and Pfaff (1984) described a small (4.0 cm diameter) probe connected to measurement electronics by fiber optics for electrical isolation. Earlier experiments along the same line were reported by Wilhelmy (1972). Misakian and his co-workers (1978) described an even smaller spherical probe (1.2 cm diameter), although this was a space-potential device in the sense that it was connected to its electronics with copper wires which had to be placed along an equipotential surface of the measured field.

These instruments were all intended for measurement of ac or transient fields. However, the work of Feser and Pfaff, and especially Misakian, has demonstrated that small probe dimensions are needed for field measurements within a few centimeters of irregularly shaped, conductive objects.

When the United States Department of Energy decided to fund a program applying fiber optics technology to power system measurements, the problem of the space-potential measurement of electric fields above ground was one that was considered worthy of study. For ac fields, the challenge was to make the probe as small as possible. For dc fields, the problem was more severe.

Only grounded measurements of the field of a transmission line had been made. Some instruments had been built that did permit above-ground fields to be measured. For example

Smiddy and Chalmers (1958) and Gathman (1968) describe large instruments built for atmospheric investigations, and quite unsuited to the investigation of the electric fields near transmission lines. Other instruments developed specifically for power line measurements (Maruvada, Dallaire and Pednault, 1983) permitted above-ground measurements, but were themselves grounded. To use these instruments, the field being measured had to be undisturbed by the presence of another object, such as an experimental animal.

For the measurement of a stationary field (eg, under a direct-current line), the induced charge on the probe must be modulated in order to make it measureable. Previously reported instruments intended for dc field measurement were designed with rotating or vibrating electrodes to modulate periodically the acquired charge. The resulting device is then known as a field mill or a vibrating plate electrometer, respectively. Comparative tests of a number of these instruments under a dc line have been reported in the literature (Comber, Kotter and McKnight, 1983). All of the instruments described operate with part of the system grounded, and are ideally mounted in the ground plane.

This section of the report describes the work done at JPL on the design of an all-dielectric ac field meter, and an isolated-probe dc field sensor. Both sensors use fiber optics, but in quite different ways. We begin with dc fields, since the work started there.

1.1.2 DC FIELDS

In contrast to the situation under an ac transmission line, ion current is a characteristic part of the environment under a high voltage direct current (HVdc) line.

The ions of interest are produced by corona processes at the conductor surface, and accelerated away from the conductor by the electric field. The density of the atmosphere is such that the subsequent motion of the ions is dominated by collisions with atmospheric molecules, and an ion of given size and charge will acquire an effective velocity determined only by the field. Such motion is usually described in terms of ion mobility. Ions may be thought of as drifting along field lines, at a velocity given by the product of the field and the mobility. Any ion current present will flow into a grounded instrument, or will deposit charge an electrically isolated instrument. As a result, the instrument must be capable of distinguishing between field-induced charge and ion current. With field mills mounted in the ground plane, the two modulated current components are in phase quadrature, and can be separated by phase sensitive detection (Maruvada, Dallaire, and Pedneault, 1983). However, the same technique, when applied to a grounded rotating cylinder field mill (Maruvada *et al*, 1983) required two sensors rotating at different speeds to separate the ion current component from the field-induced output. The device in question had two significant drawbacks. It was large (over a meter in length), and it was grounded.

The first field sensor developed at JPL was of the rotating cylindrical field mill type, but it was quite small, less than 20 cm long and 10 cm in diameter. Further, it utilized a fiber-optic readout

to isolate the probe electrically from ground. Therefore, no ion current can flow to the probe in the steady state. Instead, it will acquire a charge sufficient to prevent additional ion current from flowing to it. The effects of self-charging due to ion current are discussed briefly below.

1.1.2.1 THEORETICAL BACKGROUND

The principles involved in the design of our electric field sensor can be understood with respect to Figure 1-1.

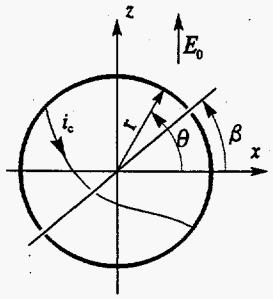


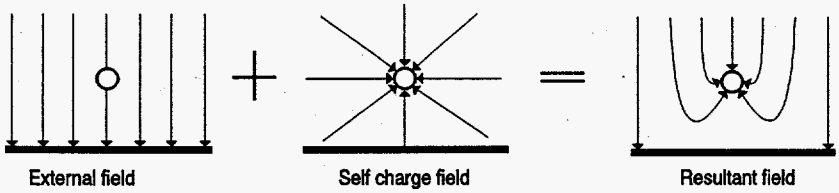
Figure 1-1. Geometry of Split-Cylinder Sensor Probe

The external electric field is determined by measurement of the displacement current between two shorted, semi-cylindrical rotating electrodes. In the figure, the radius of the cylindrical electrodes is r, and the split between the two halves is at angle β with the x axis. A uniform field E_0 is directed along z and the axis of the cylinder is along y. Under these conditions, the surface charge density σ on the cylinder at azimuthal position θ is easily derived, and is given by

$$\sigma = 2\varepsilon_0 E_0 \sin\theta \qquad (1.1)$$

The total charge on the cylinder is assumed to be zero, which means that its potential is set at the value existing at the location of its axis before it is introduced into the field.[1]

[1] It may not be immediately obvious that an object that has a potential can do so *without* acquiring a charge. It may seem that an object 1 m above ground in a 1 kV/m field, for example, must have acquired a charge *in order* that its potential be that of the surrounding space, 1 kV. The fact that this is not so can be made quantitatively clear by consideration of superposition, as in the following diagram.



In the first part of the diagram, an undistorted uniform field is shown. In the second, an undistorted radial field due to charge on the object is shown. In the third part, they are added. In this example, the sign of the object-charge is such as to cancel the applied uniform field. There is no field directly under the object, (and the field above the object happens to be double the undisturbed value). But the situation with

By integrating over one half of the cylinder, the total charge on one of the electrodes can be evaluated as a function of the electrode orientation angle $\beta = \omega t$. Differentiation with respect to time yields the current i_c in a wire connecting the two half cylinders of length l:

$$i_c = 4\varepsilon_0 E_0 l r \omega \sin\omega t \tag{1.2}$$

Therefore, if the rotating sensor is placed in a uniform field far from any conducting surface, its output is sinusoidal, with amplitude proportional to the field E_o. The current i_c is measured electronically and telemetered as a measure of the field E_o.

A similar analysis, applied to a split sphere rotating about an axis through the split, yielded a short circuit current i_s of

$$i_c = 3\pi\varepsilon_0 E_0 r^2 \omega \cos\omega t \tag{1.3}$$

This result is different by a constant factor from the cylindrical case ($i_s/i_c < 1.2$ for $l = 2r$), indicating that the finite length of the actual sensor probe does not affect the foregoing argument significantly.

If the sensor is subjected to an ion drift current, it will acquire a charge. The effect of this charge can be estimated by superposing the field from a charged conducting cylinder and the previously discussed uncharged cylinder immersed in a uniform field.

The field component resulting from the charged cylinder is independent of θ in the absence of nearby conductors, and, therefore, can make no contribution to the measured field-induced current i_c. Consequently, the sensor will not respond to a steady-state self-charge if it is far from any conducting surfaces.

The error introduced by self-charging when the sensor is near a conducting surface can be estimated using the following argument. Assume that the sensor is located in a semi-infinite region of uniform field bounded by a ground plane, with the sensor axis at distance d from the ground plane, and that an ion drift current flows toward the ground plane. The sensor, assumed to be an infinite cylinder, will intercept a small portion of the ion current, acquiring a charge Q per unit length.

Under these conditions, the field in the neighborhood of the sensor can be specified approximate-

---

no field under the object corresponds to no potential difference between the object and ground—in other words, the object distorts the field as if it were grounded.

Sine the field diagrams were drawn based on a charged object, we must conclude that *in an external field* a raised object at ground potential effectively has a charge, and a raised object at space potential (which does not distort the external field) has no charge.

ly as the superposition of three components, the uniform field (as perturbed by the presence of the uncharged conducting sensor cylinder), the field of a line charge Q on the sensor axis, and the field of an image line charge $-Q$ reflected in the ground plane at a distance $2d$ from the sensor axis. Additional secondary images resulting from the presence of the conducting cylinder in the field of the image $-Q$ are neglected.

As we have seen, the field of the line charge on the sensor axis is not detected by the sensor. However, the field component from the image charge would be detected, and gives an approximate estimate of the error if the distance d is large.

In order to estimate the magnitude of the line charge Q, note that ion current flow into the sensor will stop when the field component at its surface due to the self charge Q exactly cancels the field due to the external uniform field at points on the cylinder opposite the ground plane. When this is true, there no longer is a field component at any point on the surface of the cylinder causing an ion drift toward the cylinder. Since the field on the top surface of the cylinder at $\theta = \pi/2$ in a uniform field E_o is 2 E_o, the condition for charge collection to stop is

$$2E_0 = \frac{Q}{2\pi\varepsilon_0 r} \tag{1.4}$$

It follows that the magnitude of the field E_i at the axis of the sensor due to the image charge is

$$E_i = \frac{Q}{2\pi\varepsilon_0(2d)} \tag{1.5}$$

and the error caused by self-charging is

$$\frac{E_i}{E_0} \approx \frac{r}{d} \qquad \text{for } r/d \ll 1 \tag{1.6}$$

Again, an analogous estimate was carried out assuming a spherical sensor instead of a long cylinder, with the result:

$$\frac{E_i}{E_0} \approx \frac{3}{4}\frac{r^2}{d^2} \qquad \text{for } r/d \ll 1 \tag{1.7}$$

These two estimates set limits on the error expected for the actual sensor geometry. Note that the result does not depend on the ion current density. It assumes only that the sensor charges fully, satisfying the surface field condition. If full charging does not occur, the error will be proportionately less.

The above expressions are not accurate if r is not much smaller than d. Although exact analytical expressions in terms of a series can be derived using the image method, the slow convergence of the resultant series prompted us to find a numerical means of solving this problem. Using the finite-element technique, a numerical method was used to find the potential $V(x,y)$ everywhere

within a given enclosed region, provided that the potentials on the bounding surfaces are known. This approach was applied to the problem of an infinite cylinder near a ground plane.

Because of symmetry, only half of the structure needs to be included. This region was divided into 260 x 130 equally spaced subregions. Given V_1, V_2 and V_3 on the bounding surfaces, respectively, the ground plane, the cylinder and a plane surface parallel to the ground plane but far enough from the cylinder not to influence the results, the potential at each grid point was calculated. Knowing the voltage everywhere, the electric field can readily be calculated, and the resultant charge on the surface of the cylinder is proportional to the normal component of the electric field E_r there. Summing over all the charges on the upper half of the cylinder yields the total charge Q_u and summing over all the charges on the lower half yields Q_l. The quantity $Q_a = (Q_u - Q_l)/2$ is taken as a measure of the sensor output, since the charge distribution is not symmetrical when the sensor is near the ground plane.

The fractional change in Q_a, $\Delta Q_a/Q_{ao} = (Q_a - Q_{ao})/Q_{ao}$ is plotted in Figure 1-2 as a function of d/r. Q_{ao} is the calculated value of Q_a calculated for $d/r >> 1$, that is, with the sensor in free space.

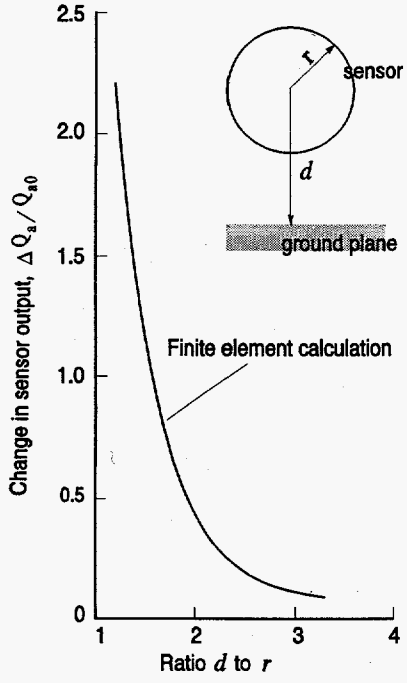


Figure 1-2. Fractional Change in Sensor Output Induced by a Nearby Conducting Surface for Full Ion Current

The quantities $\Delta Q_a/Q_{ao}$ and E_i/E_o are directly comparable, and are a measure of the change in sensor output due to the presence of the ground plane. The condition for charge equilibrium in the presence of ion current, was imposed in calculating $\Delta Q_a/Q_{ao}$.

It can be seen that a significant change in sensor output can be produced by a nearby conductor. However, these results indicate that if the sensor is kept at a distance greater than approximately five times its own radius from the nearest surface, the error will be less than 10%. In addition, estimates made using a two dimensional cylindrical model are upper limits on the perturbation

that would be seen by a sensor probe of finite length. If the sensor is not much closer to a surface than its own length, then the perturbations will be smaller than the two dimensional model predicts.

The response of similar probes in nonuniform fields has been treated by other authors (Wilhelmy, 1972 Misakian *et al*, 1978). Their results show that the sensor response indicates the field at the location of the center of the probe before the probe is introduced. Errors of the order of 1% or less were predicted when the undisturbed field varies by tens of percent over the volume of the probe. It can be concluded that the dominant error source is image fields resulting from self-charging when near a conducting surface.

1.1.2.2 SENSOR DESIGN

During the years that the work on the dc field meter was under way, two different kinds of devices were made. In the first, to rotation of the probe was done by an electric motor, powered by internal batteries. In the second, motive power was provided by the kind of turbine used in a dentists' drill, driven by compressed air. In both cases the overall instrument consisted of two parts: the sensor probe, which was self-contained and isolated from ground, and a set of readout electronics, providing an output suitable for use by a data acquisition system. The two parts were connected by conventional multimode fiber optic cable.

No particular effort was made, for the electric-drive prototype, to produce a miniature instrument. The sensor was about 8 cm in diameter and 20 cm long. The entire sensing system, including its drive motor and battery, the rotating electronics and its battery, and the fiber-optic coupler, was housed inside the split cylindrical sensing electrode pair. For obvious reasons, the device soon became known as the "beer-can" sensor. It was a proof-of-concept device, and functioned well in that role. Better measurements closer to a conducting object were clearly dependent on making a smaller probe. Therefore, considerable effort was put into developing a probe with smaller diameter. (A device with a radius of 1 cm was eventually made.)

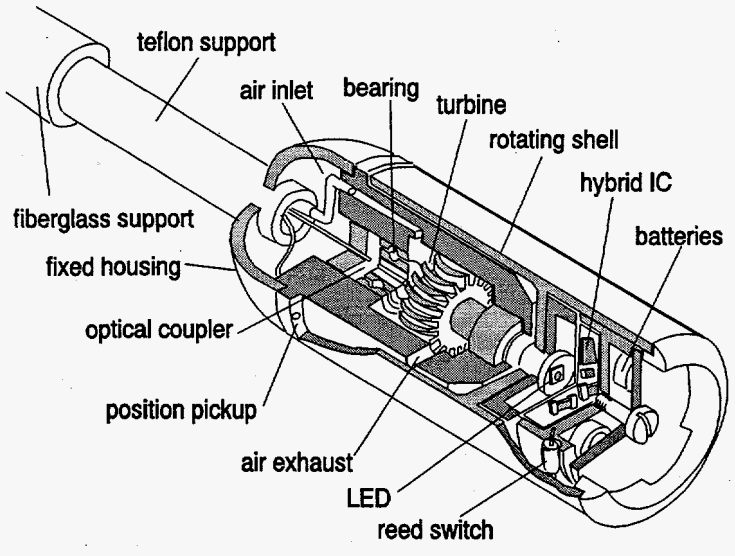


Figure 1-3. Section through dc field meter probe.

A rotary optical coupler transferred information from the rotating electronics into a non-rotating fiber cable termination mounted on the axis of the supporting tube. No use was made of slip rings or rotary transformers—this obviated a possible source of synchronous noise, and helped increase the system dynamic range.

Performance requirements were never formally determined. Unofficially, it was hoped to be able to measure dc electric fields ranging in magnitude from the normal outdoor ambient value of about 100 V/m up to the point at which the air itself breaks down, about 30 kV/cm. The accuracy goal was a few percent, with or without ions in the field being measured. It will be shown later that we came very close to achieving these performance objectives.

The first stage of the electronics converts the current at its input to an alternating voltage with magnitude proportional to the field-induced charge. The voltage modulates a voltage controlled oscillator (VCO). Narrow pulses are generated by the VCO and converted into optical pulses by a light emitting diode (LED). These pulses are injected into an optical fiber, and transmitted to the receiver portion of the system. See Figure 1-4.

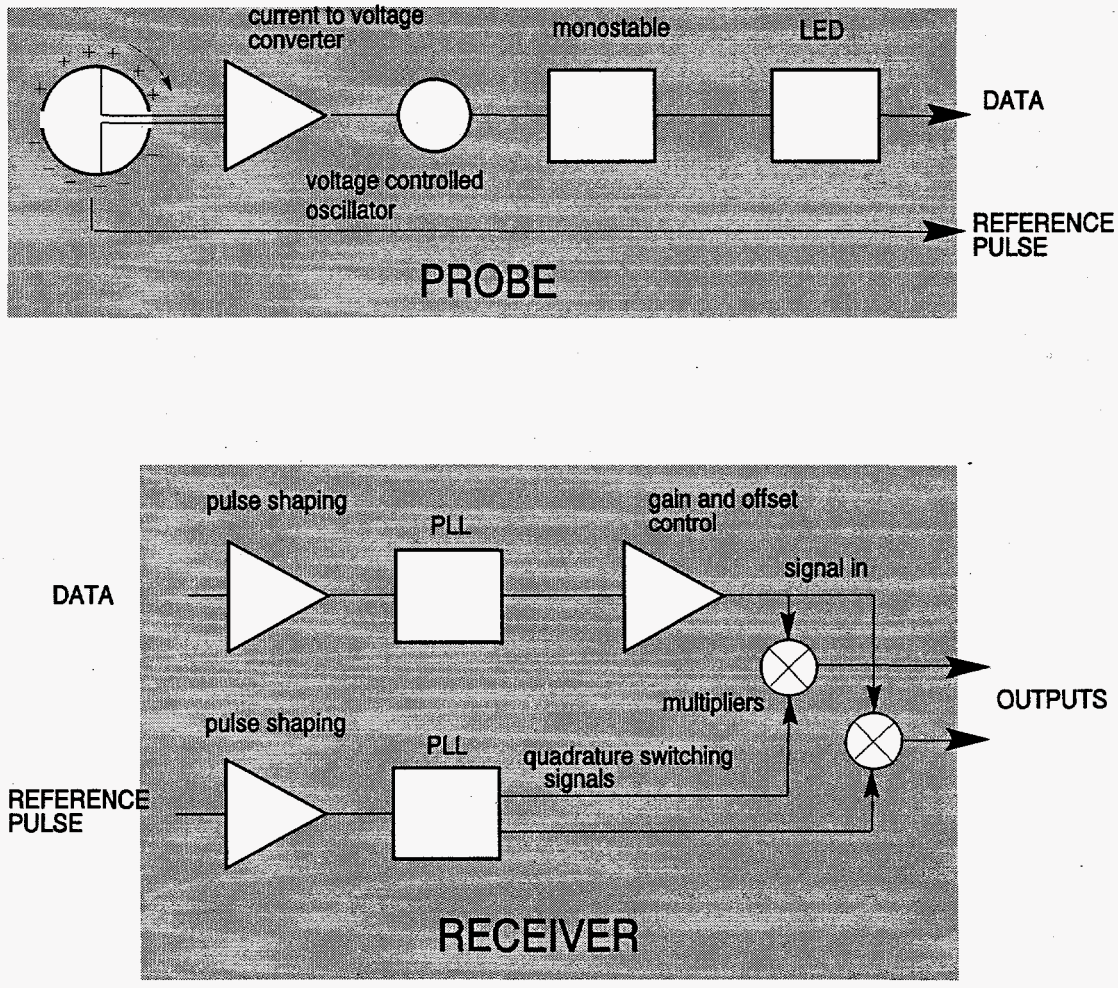


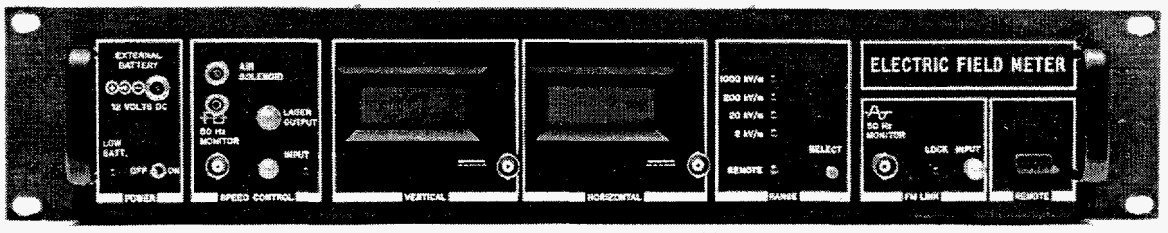
Figure 1-4. Block diagram of dc probe and receiver

To allow the easy use of a synchronous detector, an optical pulse was generated once each revolution on a separate fiber for angle reference purposes. (In the beercan version, the pulse was generated optically, and encoded onto the data pulses electronically. In that design, only one fiber was required between the sensor probe and the receiver, and the separation of the two pieces of information was electronic.) In the miniaturized probe, the pulse was generated by a small reflective spot on the edge of the rotating cylindrical shell, which reflected light back either into the fiber that illuminated it, or into a separate fiber, depending on the version. These were additional fibers to the one that carried the data. A directional coupler separated the reflected energy from the illumination in the prototypes where only one fiber was used. This lead to a much simpler design for the probe electronics, which is important in terms of reducing the size and power consumption of the hybrid IC used in the probe.

The voltage signal that modulates the probe VCO is regenerated in the receiver by an identical IC configured as a phase locked loop (PLL) detector. It is functionally similar to an fm detector in a radio. The receiver is in fact recovering the modulation from an fm carrier which has been encoded as a low-duty-cycle pulse train for optical transmission.

Synchronous detection is used to obtain maximum dynamic range, and to ease the filtering requirements in the receiver. Compared to the usual synchronous detector, such as might be used in a radio receiver, this system has the advantage of a strong synchronizing pulse, with an amplitude that is independent of the signal being detected. This pulse is used to generate the switching signals for the synchronous detector, by means of a second PLL. The switching signals for two quadrature components are generated.

In order to display the data, or to provide outputs suitable for data loggers, low-pass filtering is performed on the detected signals, as is standard practice in phase-locked detection. This effectively converts the output of the synchronous detectors into a dc voltage whose amplitude is proportional to the appropriate component, in-phase or quadrature, of the input (Johnston and Kirkham, 1989). Digital panel meters are used for data display in the prototype receiver.



Effect of pole resistance

It was observed that when the calibration electrode system (or test cage) was energized with asymmetrical voltages, the calibration of the meter was not constant. Even though the cage generated a uniform field, as verified by the field meter with the probe potential at zero, the meter indicated a higher field near the high voltage electrode. The problem was eventually traced

to leakage current down the pole supporting the probe. This current essentially grounded the probe, so that there was field intensification at the location of the measurement.

The resistance of the offending pole was independently measured (at 30 kV/m) and found to be in the range 3 to 10×10^{10} Ω/m. Thus at 30 kV, this pole could drain off nearly 0.1 mA. Although surface cleaning made a small improvement, the leakage would not go away. The leaky pole was then replaced with a fiberglass (type G-10 fiberglass, a light green material) pole. The resistance of this material was at the limit of our measurement instruments, $>10^{16}$ Ω/m. The fiber cables and air hose were also tested and found to have a resistance of $>10^{16}$ Ω/m, bringing the potential leakage current under the same conditions to $\leq 10^{-12}$ A.

The leakage phenomenon was further illustrated by observing the change in the sensor output following energization of the field cage. The results are shown in Figure 1-5 for the G-10 fiberglass pole.

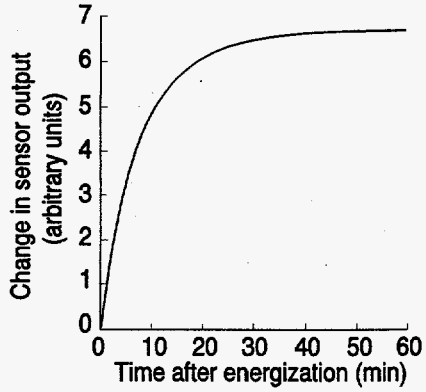


Figure 1-5. Transient change in output due to pole leakage

No ion current was present. The observed relaxation curve represents the charging of the probe-ground capacitance by charge leakage, which includes both pole leakage current and ambient ionization of the air. The time constant for the leaky pole was 20 sec, and for the G-10 pole, 13 minutes. The sensor was located 9 cm from the charged screen of the cage. The screen spacing was 1 m.

It can be seen that under these conditions, namely with the sensor near the charged conductor, supported from the ground, and with no space charge present, even the improved pole would be very unsatisfactory. Measurements might be possible with the good pole, if great care was taken to take readings quickly and then discharge the probe, but would not be possible with the leaky pole.

The addition of a small radioactive source of the type used in an antistatic camera lens brush was found to be an effective means of coupling the probe to the local space potential. The polonium-α source was attached to the base of the sensor adjacent to the rotating electrodes.

The amount of radioactivity used involves a compromise between providing an adequate source of ions to charge the sensor to equilibrium, and creating an ion cloud so dense that it perturbs

the field to be measured. Unshielded, the polonium-α source generates enough ions to measurably decrease the field seen by the probe. With the source well shielded, the leakage problem returns. To be useful, the ion production must be enough to supply the pole leakage current without changing the local field. It appears that considerable margin is available to meet these conditions.

The addition of the radioactive source resulted in constant readings after about 1 second, the instrumentation time constant. This means that the probe was remaining at space potential, in spite of the slight leakage of the pole.

The electric field data shown in Figure 1-6 were obtained with a source on the probe. The intensity of our source was estimated to be 10 mCuries. A simple calculation of the total number of ion pairs per second produced in air by a polonium-α source indicates that 0.01 to 0.1 mCi would have been adequate to drain off the expected leakage current.

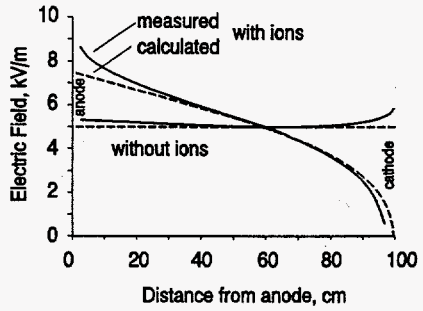


Figure 1-6. Electric Field Profile of Test Cage

The second curve in the figure was taken with a space-charge limited ion source in operation. The observed voltage versus position closely approximates the space charge limited values calculated for a planar diode (Kirkham, Johnston, Jackson and Sheu, 1987).

After some final adjustments were made (for example, the gain distribution in the synchronous detector receiver was modified), and the scale factor adjusted so that the panel meters indicated the field values directly, the linearity of the meter was evaluated.

Sometimes droop was evident below 100 V/m. On other occasions, the linearity seemed to be good down to about 10 V/m. The reason for these uncertainties is not known, but it is possible that the explanation lies in the difficulty in obtaining an accurately known field of such low value. Stray charge, accumulated on the nylon safety rope around our cage, has been observed to produce a measurable field inside the cage. Charge on the pole holding the probe could also be the culprit.

Above 100 V/m, the linearity of the probe seems to be good, perhaps in the order of $\pm 2\%$. At the highest fields measured, more than 4 orders of magnitude above the smallest readings obtained, the field in the cage was intensified so much that flashover occured. It was not clear whether or not the flashover involved the probe; in any event the probe survived.

1.1.2.3 RESULTS

The JPL dc electric field meter has been widely used in research:

> One investigator has used the meter as a hot stick to investigate a region of the electric field near to the conductor of a high voltage line, with the idea of estimating the voltage on the line from the field measurement (Feldman, 1992). His model of the field showed that, depending on the subconductor configuration, there is a region near the line where the field is not strongly dependent on the exact location of the probe. In this region, the field measurement provides a very good estimator for the line voltage.
>
> Another investigation was made of the effect of helicopter downwash on the field environment of high voltage lines. Before our measurements in Sweden (described below), there was no suspicion that the mere presence of insulators could change the field distribution. Tests of the effect of helicopters (for the purposes of evaluating live-line maintenance procedures) had consisted of anchoring the helicopter to support insulators, starting the engine, and having the pilot leave before the line was energized. Our work showed that the field distribution was likely quite unrealistic. Fortunately, our space-potential meter provided a solution: measure the field from the helicopter.
>
> A PhD student at Montana State University in Boseman used the JPL meter to investigate the interaction between snow and electric fields. He showed that the formation of snow drifts was indeed influenced by the electric charge on the falling snowflakes, and that electric charge had an effect equal to that of wind on the drifting snow. **********-***
>
> The results obtained with the JPL field meter also led to our writing contributions to an IEEE Standard. The sections on non-grounded measurements and on calibrations in IEEE Std. 1227-1990 *IEEE Guide for the Measurement of DC Electric Field Strength and Ion Related Quantities* arose from our experience.

Two sets of results in particular are worthy of detailed comment. One set resulted from visits made by the developers of the JPL meter to a test facility operated by General Electric for the Electric Power Research Institute, and the other came from a visit to the Swedish high voltage test hall of ASEA, a major worldwide electrotechnology company.

High Voltage Transmission Research Facility

The dc field meter, in various stages of development, was taken to the High Voltage Transmission Research Facility (HVTRF) in Pittsfield, Massachusetts, for measurements under their HVdc test line on two occasions. The purpose of these visits was to:

> Demonstrate measurement of dc electric fields under an HVdc line at points significantly removed from the ground potential, in a realistic environment including space charge.

Compare data obtained with the JPL meter to data from the HVTRF instrumentation.

The HVTRF is a test facility which is operated by General Electric for the Electric Power Research Institute (EPRI) for testing high voltage transmission components and techniques. It provides a short ($\approx$ 1km) section of transmission line under which our measurements were made. Two conductors were used for the dc tests, and could be energized singly or in combination. They were spaced 12.5 m (41 ft) apart and were approximately 15.5 m (50 ft) above the ground. A third conductor was located a similar distance west of these two but was not energized.

An instrumentation line was permanently installed on the ground surface along a line perpendicular to the conductors. It consisted of nine stations spaced out over a distance of 50 m (165 ft), each providing a measurement of electric field and ion current density. Electric field was monitored by a vibrating plate type of field meter mounted in an inverted position (to minimize the ingress of rain and snow) over a large plate on the ground surface. Field measurements were referred to a portable transfer standard which was calibrated in a large on-site field cage. Ion current density was monitored by a conventional one meter square guarded plate. Further detail about the instrumentation has been reported in the literature (Comber and Johnson, 1982; Anderson and Doyle, 1975) and the facility has been the site of a workshop for comparison of various types of field mills (Comber, Kotter and McKnight, 1983).

During our tests, field, ion current density and wind data were logged by the HVTRF system at 1 minute intervals, and the data were made available to us.

Field profiles across the line

Two types of measurements were made with the dc sensor in free space. One was an electric field profile across the test line, and the second involved measurements around a human subject. We will first describe the line profile data.

Profile data were taken with the line bipolar, first along a line perpendicular to the conductor with the sensor located 2.4 m (94 in) above the ground surface, and second along a vertical line midway between the conductors. The latter data were obtained with the aid of the HVTRF high-ranger (or bucket truck), and reached an extreme height of 8.5 m (28 ft). Both the vertical and horizontal components of field were read out simultaneously using the synchronous detector receiver. One exception was the point at 8.5 m, because the sensor pole was then oriented vertically, thus preventing a coordinate readout. The field was assumed horizontal at that height.

The results are plotted in Figure 1-7, viewing the plane of observation toward the north. Likely field configuration has been added (dotted lines).

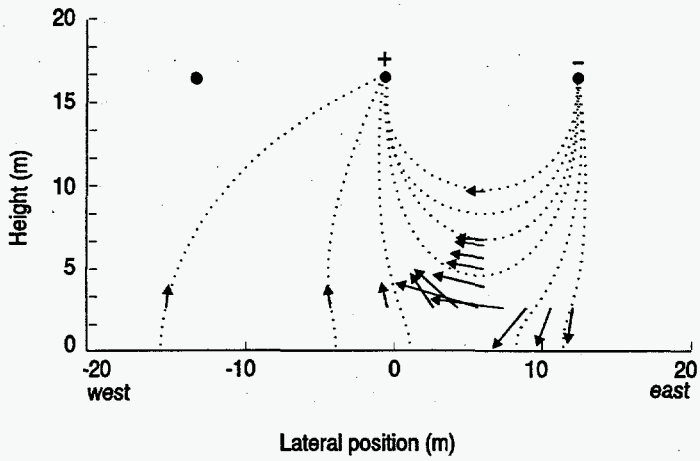


Figure 1-7. Profile of field due to line, including bucket truck data

The west conductor was +500 kV; the east at -500 kV. Rather than plot the field with lines proportional in length to the field strength, the length of each vector corresponds to the distance over which a given potential difference would occur. This makes it much easier to visualize the field. In this diagram, the scaling is arbitary. The most intense field (shortest line) was between the conductors, and was about 22 kV/m. The tail of each field vector indicates the point of measurement. The field geometry is qualitatively correct, and confirms a horizontal field at the midpoint, as is expected above the ground plane. No attempt was made to correct the measured vectors for temporal fluctuations observed by the permanent instrumentation.

One reason that this is interesting is that these results were the first ever made of the field so far above the ground.

A direct comparison was made of the readings from the JPL meter and the fixed instrumentation at HVTRF. A fairly consistent difference was found on both of our visits. Although the average difference of 11 to 15% is not large compared to the observational repeatability, it does appear to be real. The reason for the difference is not presently understood. Resources did not allow us to continue the investigation more than briefly.

Possible causes could have include one or more of the following.

1) Error in JPL sensor calibration

2) Error in HVTRF calibration

3) Local variations in field in the neighborhood of an individual HVTRF sensor

To investigate the possibility of local field variations being the culprit, a scan at a height of approximately 1 m (40 inches) was made across the instrumentation line, that is, parallel to the conductor. The data are shown in Figure 1-8.

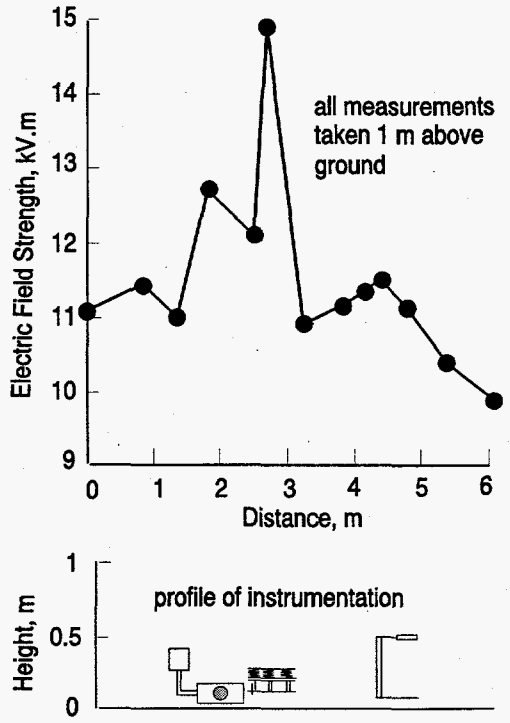


Figure 1-8. Profile across line instrumentation

The horizontal locations are approximate, being reconstructed from photographs, but the field values (ordinate) are accurately plotted. This is a repeat of a similar scan made earlier (Johnston *et al*, 1986), and again shows position dependent-variations big enough to account for the difference between the observations.

Detailed examination was not possible in the available time. The issue is discussed at length in one of the JPL reports (Kirkham, Johnston, Jackson and Sheu, 1987), where it is noted that other investigators also reported such differences. Resolution awaits further study.

While the line field profile was being measured, some localized areas of increased field were noted. The ground level field was observed to increase in spots near the test line at HVTRF. The effect was attributed to the presence, on the conductor surface, of small sources of ionization, most likely insect bodies

Most of the line had a ground-level field that agreed with calculations for a Laplace regime. The ground-level field was increased in places; next to the line itself the field was presumably reduced by the local space charge. At the ground, each tiny source mapped to a circular region of intensified field about 3 or 4 meters across.

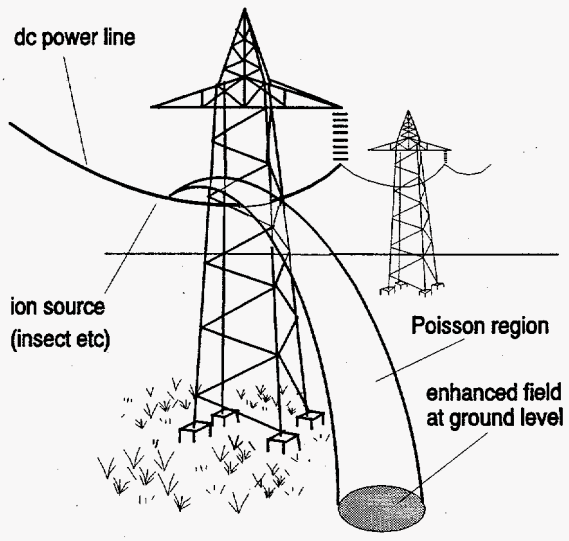


Figure 1-9. View of dc line, showing how ions created at a spot on the line surface could follow field lines and create large regions of reduced ground-level field

Measurements around a human subject

A series of observations was made to illustrate further the capability of making free-space field measurements, or field mapping in an environment containing space charge. The subject was located directly under the east conductor, and many points were mapped out in the plane bisecting his body. Both face-on and profile measurements were done, with and without a plastic hard-hat. The data, including both sensor position and field component readings were recorded photographically and reduced later. Two cameras were used, with synchronized shutters. One photographed the subject and the probe, so that their relative position was recorded. The other recorded the instrument display at the same instant. Using this approach, only 2-4 seconds were required to record one data point.

Figure 1-10 presents a sample of these field maps. The scaling used is such that the length of each vector represents a potential drop of 520 volts. The location of each measurement is at the end of the line representing the field vector closest to the subject. The inverse representation of field strength makes visualization in terms of a normal field plot of equipotentials much easier.

Interestingly, the field values over the head of the subject during our plotting efforts were observed to be 47 kV/m without a hard hat, and 27 kV/m with a hard hat, a ratio of 1.7. This number is slightly lower than the ratio of 2.2 observed on the first visit, when some spot measurements were made. The difference might be explained by a difference in resistivity between the two hard hats.

Our experience from these experiments indicates that field mapping with the dc sensor is both possible and convenient to do in the outdoor transmission line environment. The most time-consuming part of the data taking involved determination of the position of the probe, which we obtained photographically. A data logging system specifically designed for these measurements and which could automatically record an x, y, z position for the probe and E_x, Ey would be very convenient for field mapping. However, the photographic method is quite practical in the field.

For example, the field could be determined on a 5 cm grid over a one meter square plane in about 15 minutes, allowing 2 sec per grid point, a quite practical data rate. The main disadvantage of this method is that the data must be manually transmitted to a computer later.

Figure 1-10. Electric field around human subject

VISIT TO SWEDEN

From time to time, anomalous flashovers have been observed on high voltage dc insulators and bushings. These flashovers are termed anomalous because, in spite of the fact that dc insulators are bigger than ac insulators, they do sometimes flash over under normal service conditions (see, for example, Lampe, Eriksson and Peixoto, 1984). These flashovers are external to the insulator, and may cause no permanent damage, but they do result in at least a temporary shutdown of the dc system of which they are a part. No satisfactory explanation for these flashovers had been devised, and they had not been reproduced under laboratory conditions when JPL became involved.

In an attempt to understand the mechanism behind HVdc bushing flashovers, ASEA, in Sweden, scheduled a series of tests of a high-voltage dc bushing under dry and artificial rain conditions in their high-voltage hall in Ludvika. Because of the work done by the Communications & Control project on the measurement of dc electric fields in air, Task Manager Kirkham was invited to participate in these tests, and to make measurements of the electric field in the vicinity of the bushing.

This section of the Final Report describes the experimental arrangement and the results obtained.

Background Information

The term "bushing" is applied to the insulating system required to pass a high voltage conductor through a grounded object. This grounded object might be the wall of a building or the tank of a transformer. There are two general principles in the design of bushings. First, the solid insulation inside the bushing (usually oil-impregnated paper) can withstand a higher electric stress than air, and second, the likely weak point in the insulation system is the surface of the insulator (usually glazed ceramic) on the side of the bushing which is exposed to the weather. Because of these design constraints, bushings tend to be long thin objects like the insulator on an automotive spark plug and, similarly, they tend to be asymmetrical, the longer insulation being on the outdoor side.

In the case of a transformer penetration, the asymmetry is most pronounced, since the interior of the bushing is immersed in insulating oil, and the outside is exposed to the weather. The asymmetry is less pronounced in a bushing for access through the wall of the converter station in an HVdc system.

The tests performed at ASEA were of a bushing similar to the ones used on ± 600 kV systems. Altogether, this device is 15 m long, approximately 9 meters of which are on the outside of the converter station.

If we take a cross section through such a bushing, and simplify for the time being by neglecting the presence of the solid insulating material, we obtain an idea of the potential distribution, as shown in Figure 1-11. Because of the presence of the wall, the field is concentrated at the point where the conductor passes through the wall, and unless measures are taken to reduce it, the electric stress at this point may be damaging.

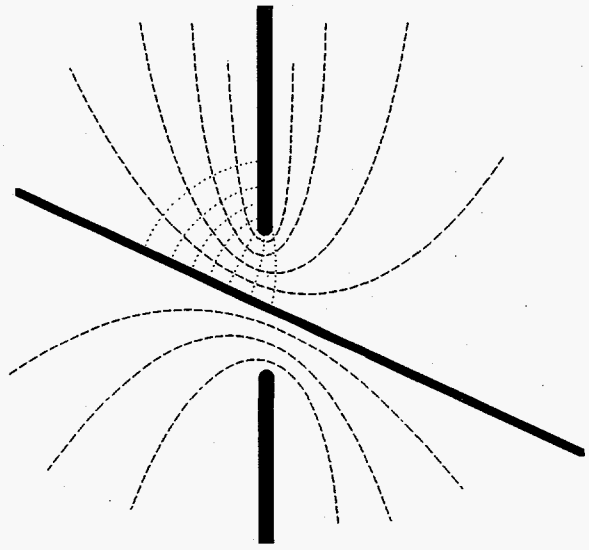


Figure 1-11. Cross sections through wall penetration in plane of bushing

If we take a cross section through the bushing across the axis of the conductor, at the point where it passes through the wall, we find that the stress is distributed non-uniformly. The highest intensity is found close to the surface of the conductor, and there is a decrease going

something like 1/r.

In order to control the stress, and to provide a measure of insurance against material inhomogeneities, conducting cylinders are inserted into the solid insulation in the vicinity of the wall, in such a way that the radial stress is more uniform, and the axial stress is reduced near the wall. In a radial sense, these conducting cylinders have the effect of making the voltage distribution more uniform. If we visualize the system as being energized with ac, it is straightforward to see that the distribution of voltage across these cylinders is governed by the capacitance between them. The cylinders are arranged so that the product of length and radius is constant. This means that capacitance between them is constant, and so the voltage drop across them is uniform.

If the bushing is energized with dc, it is supposed that the distribution of voltage goes in accordance with the very large, but finite, bulk resistance of the solid insulation material, and follows approximately the same law as that given by a capacitive distribution. (We shall see below that this idea is not supported by measurement.)

In an axial direction, these cylinders have a much greater effect. Instead of concentrating near the wall itself, the electric field stress is now spread out more uniformly down the length of the insulator. This can be seen from the equipotential lines in Figure 1-12, which is derived from a field solution calculated by an insulator manufacturer (Naito, Matsuoka, Ito and Morikawa, 1988). Note that the equipotential lines change gradually from being parallel to the high voltage conductor on the outer end to being parallel to the wall at the inner end.

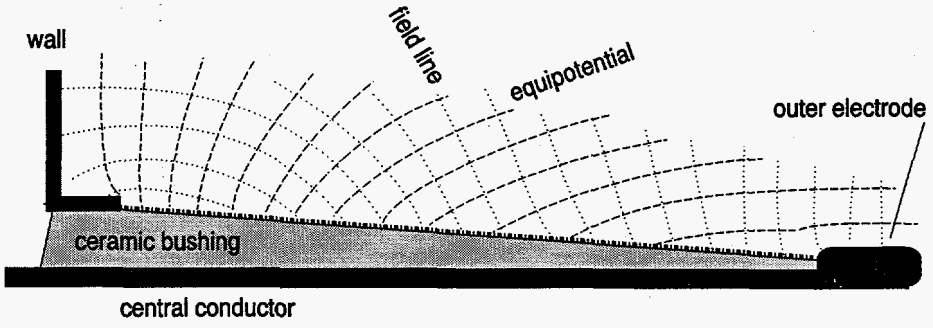


Figure 1-12. Equipotential and field lines of electric field along graded bushing

With this background, we may now proceed to a discussion of the tests performed in Sweden.

The Tests in Ludvika, Sweden

The field meter probe, mounted on a transport apparatus made by ASEA and tested earlier in Västerås, was installed near the center of the test bushing just outside where it penetrated the wall. The transport apparatus was able to transport the probe along the axis of the bushing from close to the wall out to a distance of about 2 m. The probe was installed with its axis parallel to the wall, and normal to the axis of the bushing, 50 cm from the outside of the rain sheds on

the bushing. The fiber and the air hose were about 2 m radially from the bushing. The fiber connections and the air hose then were passed through the wire mesh wall to the receiver, which was mounted at a height of about 9 m from the floor of the high voltage hall. The outputs of the receiver were connected to the recording instrumentation in the control room by way of coaxial extension cables. This indirect arrangement was used because a set of fiber optic extension cables and connectors, prepared at the last minute before leaving JPL, were excessively lossy and did not work.

The test series, which was scheduled to last approximately a week, consisted of energizing the bushing at a number of different voltages up to 800 kV, under both wet and dry conditions. In the wet tests, a very large fan was positioned so as to blow the rain away from the wall. The idea was to test the notion that when only part of the bushing was wet, the dry part experienced an increase in the electric field.

During the tests, there were a number of equipment failures. First, the positioner broke down, after a few scans of the bushing. Thereafter, it was necesary to move the field meter manually. The failure was almost certainly due to burning caused by the intense electric field. Later, the JPL meter suffered a partial failure, probably caused by a drop of water over the collimator on the position-sensing fiber. The effect of this failure was that some data were received representing the magnitude of the field, but it was not possible to determine its direction. Finally, after a couple of days, the high voltage equipment failed and the tests had to be halted.

The Effect of Incomplete Wetting of the Bushing

The ASEA engineers had theorized that, under some conditions of wind and weather, part of the bushing could be dry and part of it wet, and that this uneven distribution of wetting could lead to an enhanced electric field *in the dry zone* at the surface of the insulator, as compared to the field when the insulator was completely dry or completely wetted. Consequently, the tests with the rain and the fan on at the same time were designed to simulate conditions which would wet the insulator near the high voltage end, and would leave it dry near the wall.

Since, as we know from our tests on the fiberglass pole used to hold the probe, even a very slight leakage can look like a short circuit as far as dc fields are concerned, we would expect an enhanced field in the dry region of the bushing under these circumstances. Since the integral of this field over distance gives the voltage, the field must decrease over the wetted region of the insulator. The original intention of the tests in Ludvika was to verify these ideas by scanning the field in the intensified area.

Because of the failures of the positioning apparatus and the water drop induced problems with the position-sensing fiber on our probe during these tests, it cannot be said unequivocally that we have demonstrated that this theoretical understanding is correct. However, we are able to present evidence strongly suggestive of the correctness of this interpretation of what is happening on the surface of the bushing.

Further confirmation of ASEA's model is obtained by comparing the wet and dry results. Figure

1-13 is a composite of the data obtained from the dry tests with the positioner working, the dry tests with the positioner stuck with the probe near the wall, and the wet tests when the positioner was stuck part-way out from the wall.

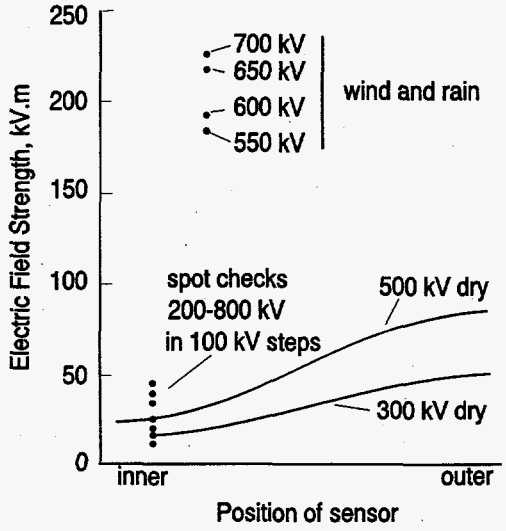


Figure 1-13. Combined results from wet and dry tests in Ludvika

It is immediately obvious in Figure 1-12 that the field close to the wall is very much higher with rain and wind applied than under dry conditions. The data for the wet conditions represent the magnitude of the dc field in the plane perpendicular to the axis of the probe, whereas the curves for dry conditions represent the magnitude of the component in the axial direction. It is assumed that the component in the radial direction is not detected, and also that there is no circumferential or azimuthal component of the field. These assumptions appear to be justified by considerations of symmetry, but they remain to be experimentally verified.

The data are self-consistent. The spot checks made when the positioner was broken (the black circles in Figure 1-12) agree very well with the dry scans at 300 kV and 500 kV made when the positioner was functioning. Also, the integral of the observed axial field under dry conditions is consistent with the voltage applied to the center conductor. The increase in field intensity in the vicinity of the wall is in broad agreement with the model proposed by ASEA for the anomalous flashovers. The increase in field in this vicinity is a factor of five when the rain and wind are applied. Unless the axial component is by far the smallest component of the field, this can only be explained by a radical redistribution of the voltage down the bushing.

The rain and wind combined cause a distribution of the voltage along the surface of the insulator which is far different than that which would be obtained under dry conditions. Later communications from ASEA indicated that even greater field enhancements were sometimes measured.

The results obtained with the field meter indicate a kind of behavior with dc that has no counterpart in ac systems. It cannot be overstressed: this kind of field enhancement would not have been widely expected. It does not happen with ac energization.

This set of experiments, and our experience with the field meter, can certainly be said to have

lead to an improved understanding of the behavior of insulation systems stressed with direct voltage. There is an important extension of our findings.

First, and most importantly, it would seem impossible to obtain a fair result by energizing an insulation system with dc as a model for ac. With ac energization, the voltage distribution in a large insulation system is governed by capacitance. With dc energization, the voltage distribution is governed by surface conditions. Only if bulk (volume) effects were dominant would dc energization produce a field representative of ac energization. Bulk effects do not dominate: the quality of solid insulation is so high that volume-leakage is negligible.

This means that it is unfair to estimate the ac insulation performance of a bucket truck tested with dc. DC testing is used fairly routinely as it is much easier to detect partial discharges (PD) with dc energization. With ac applied, there is a large capacitive current that makes PD measurement difficult. We have now strong evidence that dc conditions depend on surface effects, and ac on capacitance. While we have no direct evidence, it is conceivable that a large object (such as a bucket truck) would pass one kind of test and fail another.

Since bucket trucks are tested for safety in live-line maintenance, it would seem prudent to investigate this question further. We will gladly make our field meter available for any such tests.

1.1.3 AC FIELDS

1.1.3.1 INTRODUCTION

The problems of measuring ac electric fields are far easier than those of making the same measurement in an dc field. In fact, devices to make space-potential ac electric field measurements have existed for some while. One such instrument which has been used in the power industry is described by Deno and Zaffanella (1975). In this device, the sensing probe and the readout are integrated into one housing, which is held beyond arms length at the end of a fiberglass pole during operation. It was the existence of this meter, and others like it, that pointed up the non-availability of such instrumentation for dc fields.

The capabilities and features of the existing ac field meters provided, in a way, the baseline performance parameters for the DOE dc field meter system. Since an object of the dc field meter development was to put the measurement of dc fields on an equal footing with ac fields, as the dc meter was being evaluated, comparisons of its performance with the performance of similar

instruments designed for ac use seemed natural. While the dc field meter is somewhat more complicated, it provided performance advantages in several areas over its ac counterpart:

Size: The dc probe was smaller than most commonly used ac probes. This means that readings could be obtained of fields that were uniform in extent over quite small regions. The unit could be used close to distorting objects.

Dynamic Range: The dynamic range of the dc probe was so large that range switching on the probe was unnecessary. This means that there was no need to know the approximate value of the field before measuring it. In addition, the probe could be moved from a low field region into a high field region without the need to approach it.

Power: Some versions of the probe had no externally controllable power on/off switch, being switched by a centrifugal switch. There was thus no need to approach the probe to turn it on prior to a measurement, or to turn it off afterwards.

Data Acquisition: The JPL dc field meter could be readily connected to a data acquisition system. This was not always true of space potential instruments.

While some of these improvements in performance arose specifically from improvements made to solve the problem of dc measurements, it seemed natural to ask whether some of these advantages could be retained in a redesign of the system for use in ac fields.

Size: The small size of the dc probe was made possible because of the use of hybrid integrated circuit techniques, and because of the small size of the drive motor (turbine). An ac probe could use similar electronics, and would not need a drive motor.

Dynamic Range: The large dynamic range of the dc probe was largely a result of the use of coherent (synchronous) detection. This approach could also be implemented in the case of an ac probe, even though a synchronous reference pulse was not readily available.

Power: The lack of power switch of the dc probe was a response to the size goals of the design, and took advantage of the rotation of the sensor head. This option would not be available in an ac probe.

Data Acquisition: The dc meter could be connected to a data acquisition system because part of the meter system was at ground potential, interconnected with the probe by means of fiber optics. This design feature could readily be retained in an ac design.

It was therefore undertaken to design an ac field meter that retained as many of the advantages of the dc meter as possible. Two approaches were followed. First, a completely passive design based on the electrometer was developed. Later, an electronic version, inheriting much from the dc meter, was produced.

1.1.3.2 ELECTROMETER

A number of feasibility models were fabricated, and test data were obtained. The basic principle involved was to detect a field-induced force on a quartz fiber or fibers by measuring their flexure optically. The fibers used in the electroscope were optical fibers with core and cladding, which could readily be connected to a monitoring instrument via cables containing similar fibers. The basic idea was first described by Kirkham, Johnston, Lutes, Daud, and Hyland, 1984.

There are a number of attributes, both good and bad, that distinguish the quartz fiber electroscope from other devices for field measurement. On the good side, the device could be made very small: a package on the order of 5 mm in diameter and 20 mm long should be quite practicable. No power is required at the device, it is completely passive. It is possible, as will be described later, to build the sensor of dielectric materials only. No conductors need be present. This is attractive in applications involving measurement of very high fields where corona effects or air breakdown at structural discontinuities must be guarded against.

On the negative side, the device could be predicted to be non-linear, and to be sensitive to acceleration. The latter point, acceleration or g-sensitivity, is strongly dependent on the electroscope geometry. A single cantilevered fiber used as the sensitive element is clearly g-sensitive. Experimental tests showed that 1g could produce an output corresponding to a significant fraction of the full scale deflection. On the other hand, it is possible to compensate for acceleration-induced response by building a two-fiber electroscope such that both fibers deflect together, eliminating significant relative motion under g-forces, but deflect oppositely due to electric field-induced forces. Compensation of g-forces or vibration effects were not investigated experimentally. A final factor that is not necessarily emphasized is that the device responds inherently to E^2, which makes it blind to the sign of a field. It will measure fields in either direction along its sensitive direction, consistent with its being considered an ac sensor. (In fact, the device can measure high frequency or RF fields as well as those at power frequency, which may suggest other applications.)

In order to facilitate building of test models of the electroscope by a skilled technician, but one untrained in techniques for making fused silica, a design was worked out which was based on building the device up from microscope slides and glass spacers cut from microscope slides. UV curing cement was used for assembly.

The configuration is illustrated in Figure 1-14, which shows an all-dielectric implementation. In another implementation, a conductive coating was placed on the fiber itself and on a very small area of the mounting slide immediately under the fiber. In both case the sensitive field direction is perpendicular to the microscope slide base.

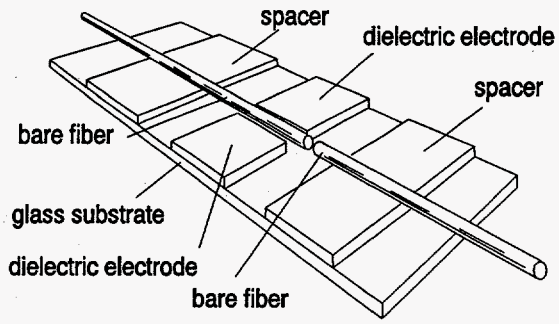


Figure 1-14. All-dielectric version of electrometer field meter

In view of the nonlinearity of the device (see Figure 1-15), and the difficulty experienced in aligning the two fibers, we did not develop the system extensively.

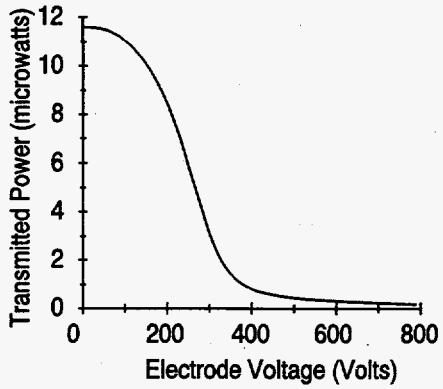


Figure 1-15. Input/output curve of electrometer field meter

A NASA Tech-Brief was written about this device. Somewhat to our surprise, we found out a year or so later that Edjewise, a company in the midwest, had produced and sold fault indicating devices based on the Tech Brief approach for use in distribution systems.

A number of applications for these unique all-dielectric sensors were discussed with other workers in the power industry. Examples are briefly described below.

(1) It should be possible to monitor the state of very large circuit breakers using variations in field near the gap. This field is dependent on the position of the contacts. It would not be possible to directly measure contact position with any conventional devices because of the potentials involved. This might provide a mechanism to detect "stuck breaker" conditions by external measurements, independently of the breaker auxillary contacts.

(2) A *hot-stick* with a very simple and robust sensor encapsulated in it would be a valuable safety tool for service personnel. The sensor could detect a live conductor on approach and trigger an alarm signal. These JPL sensors are particularly suited to this application since they are very light.

(3) A simple dielectric device that could be manipulated in proximity to a long insulator

string could be used to detect a shorted module in the string. Multiple failures of this type could result in a failure of the entire string. Failure (puncture) detection of this kind could improve insulator string performance.

A simple test of this idea was conducted at General Electric's High Voltage Transmission Research Facility (HVTRF) in Pittsfield, Massachusetts during a visit there by JPL personnel. Detection of a deliberately shorted element in a short insulator was readily done.

Further development of this approach was not pursued at JPL. The work there concentrated on the ac equivalent of the dc field meter, described next.

1.1.3.3 ELECTRONIC FIELD METER

The shape of most existing field meters has been governed by the need to include the readout device in the probe. Accordingly, the side of the probe facing the operator was flat, and the probe overall was box-shaped. Since the proposed ac field meter did not need to rotate, and the readout would be at the remote end of a fiber optic cable, its shape could be chosen arbitrarily.

A spherical shape has two advantages: its sensitivity can be calculated from theoretical considerations, and it results in minimum disturbance of the field. Therefore, this shape was chosen for the DOE ac field meter probe.

Spherical field meters have been built, but have not been used to any significant extent in the power industry. A small spherical probe was built by Misakian, Kotter and Kahler (1978) at the National Bureau of Standards. This probe was only 1.2 cm in diameter, and contained no electronics. The system was designed for use in the uniform field of a test cage. However, it was not truly a space-potential device since the probe was connected to its battery-operated electronics by means of a pair of conductors, and space potential was guaranteed by adjustment of an external circuit. In the test cage for which it was designed, the equipotential forced by the output wires was unimportant: the wires could be arranged to be along known equipotential lines.

More recently, in Germany, Feser and Pfaff (1984) built a spherical probe for the measurement of transient electric fields. This probe was similar in some respects to the one we proposed to build; however, the device was quite large (4 cm diameter) perhaps because it was designed to give the x- and y-components of a transient electric field. A somewhat simpler probe was built by Gorakhpurwalla, Cooper and Johnson (1984). This device contained battery-operated electronics, and was 4 in (10.16 cm) in diameter. A different approach was taken by Friedmann, Curzon, Feeley, Young and Auchinleck (1982). Their probe consisted of a gas-filled bulb 2.5 cm in diameter, and contained no metal parts. The device had the interesting property that it had no directional sensitivity. Its dynamic range was rather small, however.

The size of the JPL ac field meter probe was set by the fact that a suitable hybrid IC had already

been designed for the dc probe. A hollow spherical probe of only 1 cm radius (2 cm diameter) would be less than half filled by the existing hybrid IC. This choice would leave a space of equal size for the power source. While it seemed possible to make an even smaller probe, the extra work involved did not seem justified.

As far as possible, the design of the ac field meter was based on that of the dc meter. The probe used the same IC, with some slightly different component values, and the receiver used a similar coherent detection method. The method of energizing the probe was, however, quite different.

Power Source

It was intended from the outset to use a photoelectric source for the ac field meter probe. The approach is discussed by Kirkham *et al.* (1984) and work done on the problem is described by Johnston *et al.* (1986). The system consists of a laser source at the base station whose output is coupled to a fiber. At the other (sensor) end of the fiber, the light is coupled to a photovoltaic cell or array to produce an electrical output. Figure 1-16 shows the proposed means of optically energizing the probe electronics, and the losses that are associated with each step.

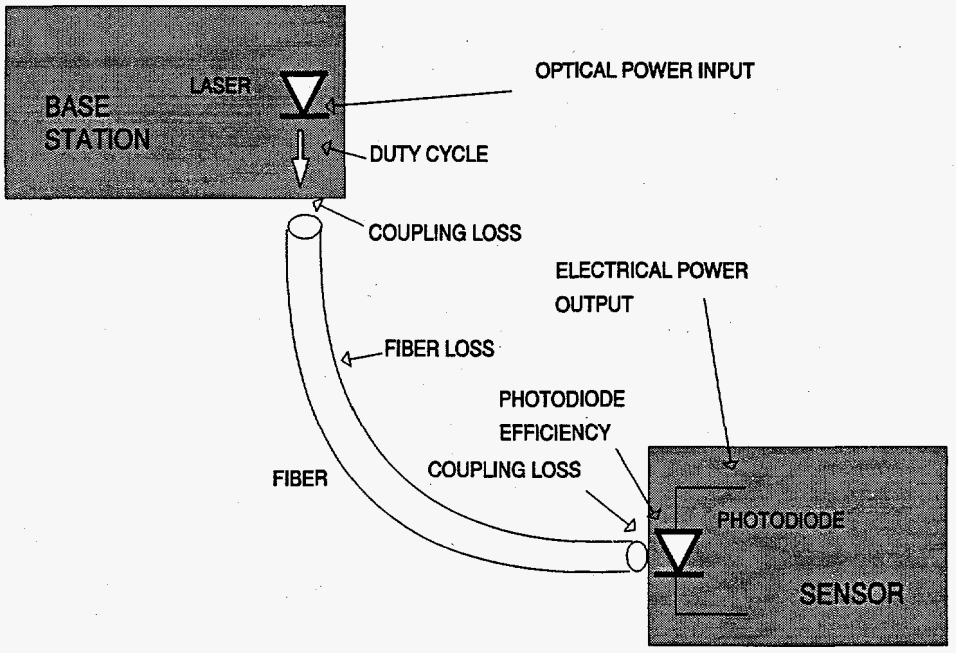


Figure 1-16. Block diagram of optical power transfer scheme, showing loss mechanisms

Integrated-circuit photodiode array

A small integrated-circuit array of photodiodes was designed to power the ac probe. While production problems eventually forced us to use another approach for the prototype probe, several samples of the IC array were made. (Johnston and Kirkham, 1989).

Fabrication of the photodiode array starts with the epitaxial growth of three layers on a semi-insulating (SI) gallium arsenide (GaAs) substrate. The SI substrate is required because of the

need for series connection of the photodiodes. Silicon has no semi-insulating state, and therefore the production of silicon photodiode arrays cannot be done using only integrated circuit techniques. However, inexpensive arrays can nevertheless be manufactured, and several were used in our optically powered work.

Gallium arsenide arrays, made using IC fabrication methods are also available and, while they are more costly, they are also more efficient. This topic is reviewed in the next section of this Final Report.

Mechanical Design

The DOE ac field meter probe is mechanically very straightforward, as shown in Figure 1-17.

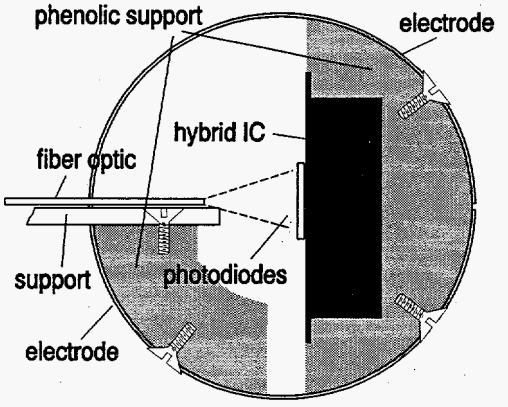


Figure 1-17. Cross section through the ac field meter probe

The outer shell of the probe forms the electrode pair, in much the same way that the shell of the dc probe comprises the semi-cylindrical electrodes of that device. In the case of the ac meter, the electrodes are spherical. Since there is no motion associated with the ac meter, there is no need to consider the weight of the structure or its balance. Spun brass hemispheres, with an outside diameter of 2 cm, form a relatively simple and sturdy housing for the device.

Inside the sphere, the various parts are held in the proper relative position by two pieces of machined phenolic material. One piece holds the hybrid integrated circuit, and insulates it from the electrodes. Both electrodes are attached to this piece. The other phenolic part is attached to one of the electrodes, and is used to hold the entire assembly to the handle.

Three optical fibers are used. Two bring power to the electronics, via gallium arsenide photodiode arrays. There is an air gap of about 4 mm between the ends of the fibers and the diode arrays, to allow the incident power to spread over the entire array. The other fiber has an LED attached to it, and carries the data from the probe to the receiver. The three fibers are held against a fiberglass "paddle" as they enter the probe. This provides a measure of mechanical support for the fibers.

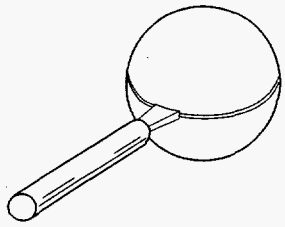


Figure 1-18. AC field meter probe, actual size

There is a small rectangular opening in the shell to admit the paddle. The opening is about 1 cm wide and 1 mm high. The "missing" electrode area is therefore about 0.1 cm^2. Since the total surface area of the electrodes is over 12.5 cm^2, the error caused by the opening is not likely to be significant.

Electrical Design

The ac field meter was designed to take advantage of the work already done for the dc meter. Accordingly, the two are very similar electronically.

Probe Design

The probe electronics for the ac probe is based on the same data encoding scheme as the dc probe. A charge measurement circuit modulates a VCO, which produces short optical pulses at a frequency proportional to the input current. The configuration of the hybrid is identical to the hybrid used in the dc meter; the values of some of the components are changed because of the different size of the electrodes, and the lower frequency of the input.

Additional changes were made to accommodate the limited power available from the optical power source. A low-power amplifier was used in the first stage of the probe, and the frequency of the VCO was decreased. In the prototype system, about half the power of the IC was dissipated in the first stage. This power was the quiescent power of the operational amplifier used. When this was reduced by the use of a different amplifier, the dominant power user in the hybrid became the LED transmitter.

With optical pulses as narrow as practical from the hybrid circuit already made ($\approx$100 ns), the power was further reduced by decreasing the free running frequency of the VCO to 10 kHz. This was 9 times lower than the frequency of the VCO in the probe of the dc probe, and further reduced the power consumption. The first miniaturized dc probe had consumed 8.4 mW—the power demand of the final ac probe was less than 2 mW.

Active "zener" diodes were used to stabilize the voltage applied to the hybrid IC. There are small power variations caused by changes in the fiber loss as the fiber is bent during normal use. It was found that, without stabilization, these power fluctuations resulted in large changes in the power supply voltage, and modulation of the VCO. While this modulation is not at power frequency, it did add low frequency noise to the signal. The lock range was therefore reduced at small signal. The zener diodes completely eliminated this effect.

Receiver design

Since the data encoding method is the same as that used for the dc meter, it may be expected that the receiver design is similar, and indeed, such is the case. There are two differences. They arise from the fact that the dc meter has a synchronous reference pulse obtained from the rotating shell of the probe, and there is no corresponding pulse in the case of the ac meter. This reference pulse enables the dc meter to use coherent detection in a very straightforward way, and to present the data resolved into two components.

Without the reference pulse, coherent detection becomes more difficult, and the output cannot be resolved into components. During development, envelope detection (rather than synchronous detection) was evaluated for this application, in view of its simplicity. The dynamic range was found to be much less than that of the dc meter. It became clear that synchronous detection would have to be used to achieve the potential dynamic range of the instrument.

One of the early uses of the phase locked loop, before it became available on an integrated circuit, was the tracking filter. In particular, the PLL was applied to the problem of recovering signals deeply embedded in noise. Much of this work was done at the Jet Propulsion Laboratory (see, for example, Gilchriest, 1964)[2] where the object was to recover the weak signals from distant spacecraft. Apart from the much lower frequencies involved, our present problem is the same.

It is convenient to view the problem in terms of a radio receiver. The system consists of a frequency modulated carrier where the modulation happens to be at a fixed frequency (60 Hz, in the case of the ac meter). It is the amplitude of the modulation signal (ie, the magnitude of the 60 Hz signal) that is of interest. A phase locked loop is readily used for the FM demodulation, because the FM carrier is very strong (the 10 kHz optical pulse train). Unfortunately, the process of frequency modulating the carrier, and particularly the process of demodulating it, results in the addition of a large amount of noise to the recovered modulation.

This is particularly true of the kind of data link used in the field meter. With a large modulating signal, the carrier frequency can be significantly different from its nominal value, and may even approach zero. No demodulating scheme could be expected to reproduce the modulation when the carrier frequency is so low. Distortion caused by very low carrier frequency has been observed on the recovered signal at large fields. It is, nevertheless, a straight-forward matter to recover the signal.

At the other extreme, with a very small signal, the noise of the carrier demodulation process may be larger than the signal itself. In the limit, even with no modulation, there will be noise at the FM detector. The problem of recovering a signal from this noise presents more challenge.

---

[2]Coincidentally, Carl Gilchriest was, during the early 1980s, the Manager of the JPL Project that later did the work reported here.

In a manner analogous to the application of a PLL to the problem of reception of weak AM signals, it is possible to use phase lock to measure the amplitude. Because the error voltage in the loop can be extensively filtered, a clean switching signal can be generated by phase locking an oscillator (at the modulation frequency of 60 Hz) to this (noisy) input signal. Only loop response speed is lost as the filtering becomes more extensive. The frequency of the signal, ie, the power system, is unlikely to change appreciably, either in the short term or the long term.

The switching signals for the ac field meter are derived by means of a phase locked loop from the signal itself, whereas in the dc meter, these signals were derived from the synchronous reference pulse. Figure 1-19 compares the receivers of the dc meter and the ac meter.

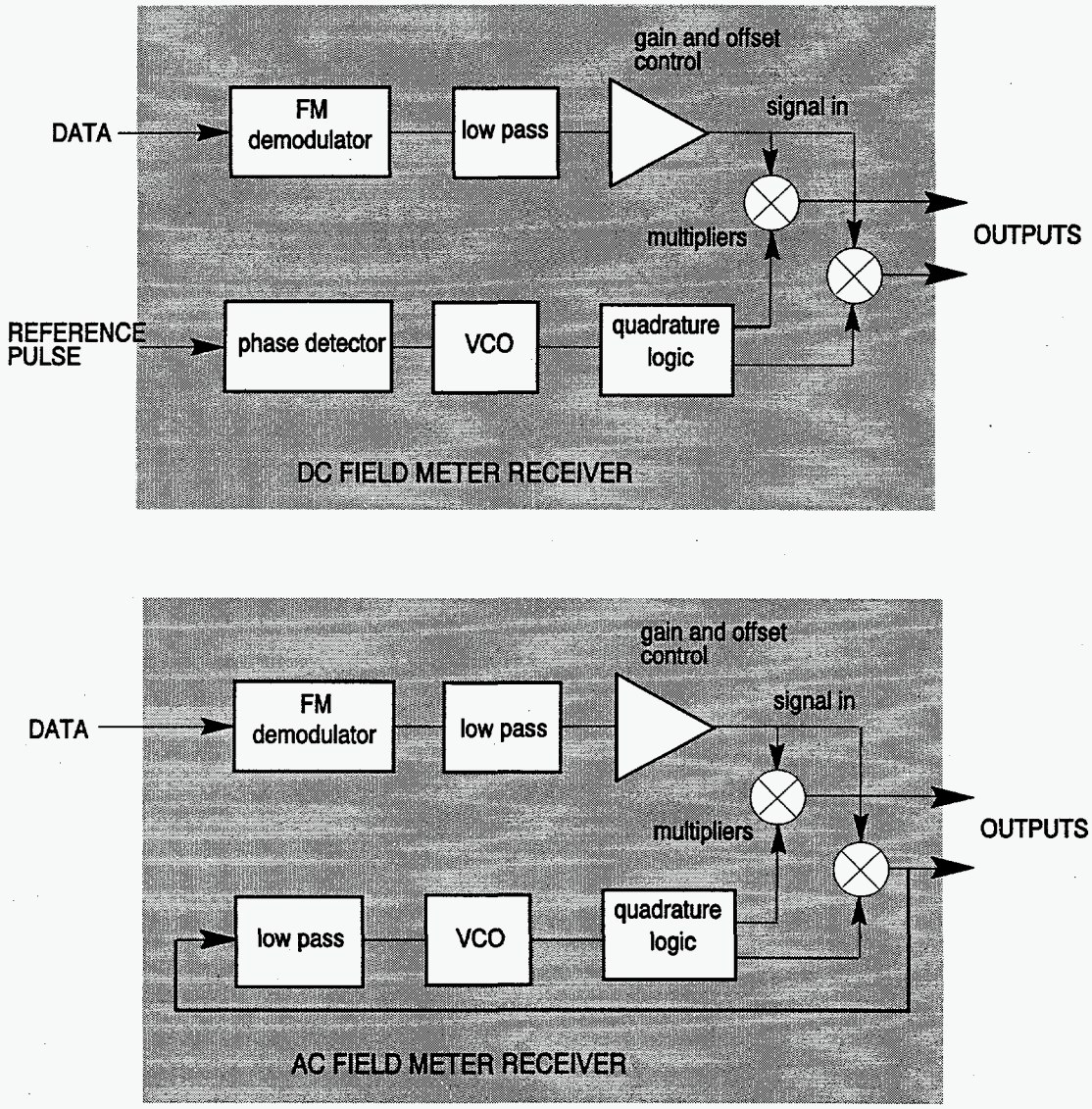


Figure 1-19. Comparison of ac and dc field meters

A small number of changes in the printed circuit for the dc field meter converted it into one suitable for use with the ac meter. Components were changed because of the lower FM frequency. The PLL used to generate the quadrature switching signals, originally in what was described as the reference channel, was modified to have a free running frequency of 4×60 Hz. One of the two synchronous switches was used as the phase detector for this loop, so that the loop locked to the signal rather than the external reference.[1] The panel meter originally used to display the magnitude of the signal component was retained as a lock indicator. In the time since the dc meter was developed, an improved synchronous detector was designed.

Passive filtering was used to drive the panel meters, without further amplification. The receiver gain distribution was adjusted so that there was no range changing after the synchronous switch. This, and the use of high-quality amplifiers in the conditional inverter stage, eliminated the need for an output offset adjustment.

1.1.3.4 CALIBRATION

Because the ac field meter is spherical, the distortion it causes in an otherwise uniform field can be calculated. This results in the possibility of a calibration from first principles, rather than an experimental calibration. Wilhelmy (1973) examined the distribution of charge around a conducting sphere in a time-varying homogeneous field. His work was aimed at a probe that could provide information on the x-, y-, and z-components of the field. Of more direct application to our situation, Deno and Zaffanella (1975) derived the expression

$$i = j\omega 3\pi r^2 \varepsilon_0 E \tag{1.8}$$

where i is the current between two hemispheres whose polar axis is aligned with a field of magnitude E. The radius of the sphere is r, ϖ_0 is the permittivity of free space, and ω the angular frequency. To obtain a calibration from first principles, all that is needed is to inject a known current at power frequency and adjust the gain of the receiver until the display reads the value of field corresponding to the input current.

A high quality calibrated signal source was used to generate power frequency current. The signal source was terminated with a 50 Ω load, and capacitors were used to feed a small amount of current to each of the electrodes in the probe. The necessary adjustment was then made in the receiver. The arrangement is shown in Figure 1-20.

[1] It is usual, in the design of radio receivers, to have two phase detectors in the PLL system. One acts as the synchronous detector, and the other generates the loop error signal, as here. The two phase detectors are usually operated in phase quadrature by phase shifting the signal 90 degrees. Since the logic to generate quadrature switching signals (rather than quadrature versions of the received signal) had already been developed for the dc meter, that approach was retained here.

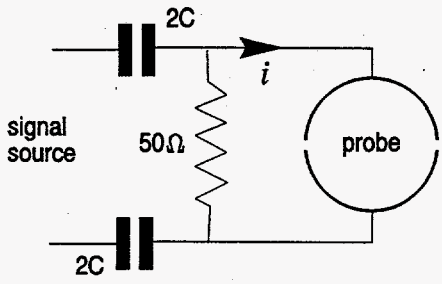


Figure 1-20. Calibration arrangement for ac meter

The injected current, i, in Figure 1-20 is given by

$$i = V_{app} \times \omega C \tag{1.9}$$

where C is the value of the two capacitors in series, and V_{app} the applied voltage. The current can be eliminated between the previous equations:

$$E = \frac{V_{app} C}{r^2} \times \frac{1}{3\pi\varepsilon_0} \tag{1.10}$$

With the values used, equation 2-3 gives that $E = V_{app} \times 63.308$. If V_{app} is measured in mV, the field is expressed in V/m.

Sources of uncertainty in calibration

When it was inserted into an electric field, the probe measured a value that seemed to be a few percent different from the one calculated for that field. The discrepancy was eventually traced to distortion of the waveform of the ac source energizing the test cage. Once this distortion had been removed, the probe seemed to measure very accurately the value of the field.

This high degree of accuracy should not be surprising. Most of the sources of error in the receiver are removed by the method of calibration. Remaining sources of uncertainty in the calibration include uncertainty in the value of injected current, errors in the range-changing resistors in the receiver, and digitization error.

With a 3½ digit display, the digitization error can normally be held below one percent. On the range for which calibration was done, the predominant uncertainty is in the current fed into the probe. The measurement uncertainty is the sum of that due to the current, and that due to the range change resistors. One percent resistors are used in the range-changing switch in the prototype receiver.

As mentioned above, the current injection was from a high quality signal generator terminated in 50 Ω Series capacitors from the 50 Ω termination to the probe provided essentially a current source. It is difficult to measure such a small current; however, the current is eliminated in the procedure we used. Instead, it is necessary to know the voltage across the 50 Ω terminating resistor and the value of the series capacitors.

While the voltage across the resistor could be estimated from the signal generator readings, it can also be measured as an AC voltage to an accuracy of better than 0.5%. Since the injected current is given by this voltage divided by the impedance of the series capacitors, uncertainty in the value of the capacitors adds to the uncertainty in the measurement of voltage to produce the total uncertainty in the current. In our calibration, 10% capacitors were used. This 10% would become the dominant uncertainty in the calibration procedure. To reduce this uncertainty to an acceptable value, the capacitors were measured. The two 1-nF capacitors in series measured 527 $\pm$3 pF.

One remaining source of uncertainty is the precision with which equation 2-3 above fits the geometry of the probe. If the probe is not perfectly spherical and divided precisely in two, the equation cannot be expected to fit exactly.

The housing of the probe was fabricated from spun brass. It is expected that each hemisphere is identical to all the others (20 were made in a single batch). Consequently, no attempt was made to measure how precisely the sphere was bisected. Measurements of the hemispheres used in the probe indicate that the average radius is accurate to 0.2%, but that the probe is spherical only to within about 0.5%. An average weighted to enhance the contribution of the polar direction of the sphere was used for the radius in equation 2-3. It should be noted that the missing electrode area (necessary to accommodate entry of the fibers and the mounting paddle) is also very small, accounting for less than 1% the total surface area. Further, this surface area is at a part where, during operation, the induced charge is close to zero. No appreciable error is therefore expected from geometrical considerations of the probe.

Considering all the above uncertainties, the ac meter can be expected to have an uncertainty of about 2% on any range. The accuracy should be somewhat better on the range for which calibration was performed.

1.1.3.5 LABORATORY TESTS

A number of tests were done to examine the performance of the ac field meter. Of particular interest was the question of the dynamic range, since the phase-locked loop detector was different from the one used in the dc meter; the stability of the probe, with its optical power supply was also of interest.

Later, the meter was taken to the National Bureau of Standards now the National Institute for Standards and Technology), in Gaithersburg, Maryland, for tests in a field known to greater accuracy than any we could establish at JPL.

Dynamic range

The considerations that determine the dynamic range of the ac field meter are not the same as for the dc meter. In the case of the dc meter, the upper limit of signal-handling capability is

defined by saturation of the signal chain and the lower limit by the level of synchronous noise in the system. In the ac meter, the upper limit is the same, but the lower limit may be set by other effects. For example, because the phase-locked loop used to synchronize the detector to the signal operates on the signal itself, and not a reference pulse, the loop can lose lock when the input signal is small or changes abruptly. This cannot happen in the dc meter.

Theoretically, the receiver can measure fields that produce voltages far below the level of random noise in the signal path. In practice, this limit has not been achievable because of the loss-of-lock problem. In addition, the 240 Hz oscillator used to produce the synchronous switching signals is subject to drift, and consequently there is a phase error when the field being measured is small. With synchronous detection, this phase error translates to an amplitude error.

Demodulator noise begins to mask the signal at about 50 V/m. Although lock can be maintained for signal levels 20 dB lower than this, a practical lower limit for the prototype meter is about 10 V/m, with an upper limit of 2 kV/cm.

1.1.3.6 TESTS AT THE NATIONAL BUREAU OF STANDARDS

On 8 April 1987, the ac field meter was taken to the National Bureau of Standards to examine the performance of the meter in a field known more accurately than any we could produce at JPL. The field was produced by applying a known voltage to an accurately measured parallel-plate arrangement. The field meter had been calibrated from first principles. The results of the tests are shown graphically in Figure 1-21, which is drawn so as to exaggerate the differences between the calculated and measured values of the field.

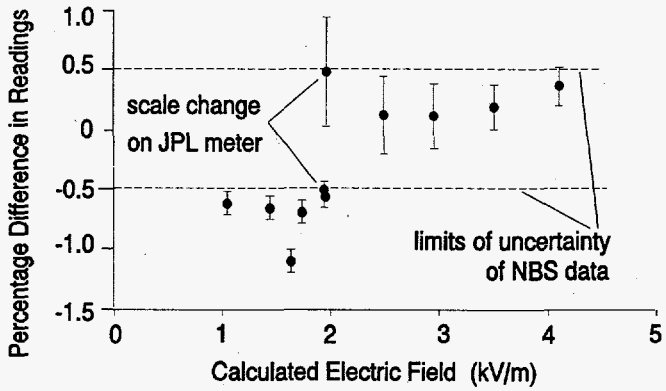


Figure 1-21. Differences between calculated and measured fields in NBS tests

Although a 5-digit meter was used to measure the voltage applied to the plates, it was the opinion of Dr. Misakian, who conducted the tests, that the uncertainties in the system might total 0.5%. In the figure, horizontal dashed lines are drawn at ±0.5%. It can be seen that the measured values of the field are almost all within or close to the 0.5% lines.

The exception is the point corresponding to a field of 1.66 kV/m. It happens that this was the first point measured. There is a possibility that something in the receiver or probe electronics

had not, at the time the reading was taken, reached a stable operating condition. If allowance is made for the digitization uncertainty (vertical bars in Figure 1-21), the readings are within 1% ± 1 digit.

The change in sign of the differences at a field value of 2 kV/m occurs because of a range change on the JPL meter. (The meter had been calibrated on the least sensitive of the two ranges—on the right in the figure—used in the tests at NBS.) A trim resistor on each range would reduce this effect, but such a refinement is not thought worthwhile.

An indication of the linearity of the meter over a wide range may be obtained from Figure 1-22, which shows calibration data obtained synthetically.

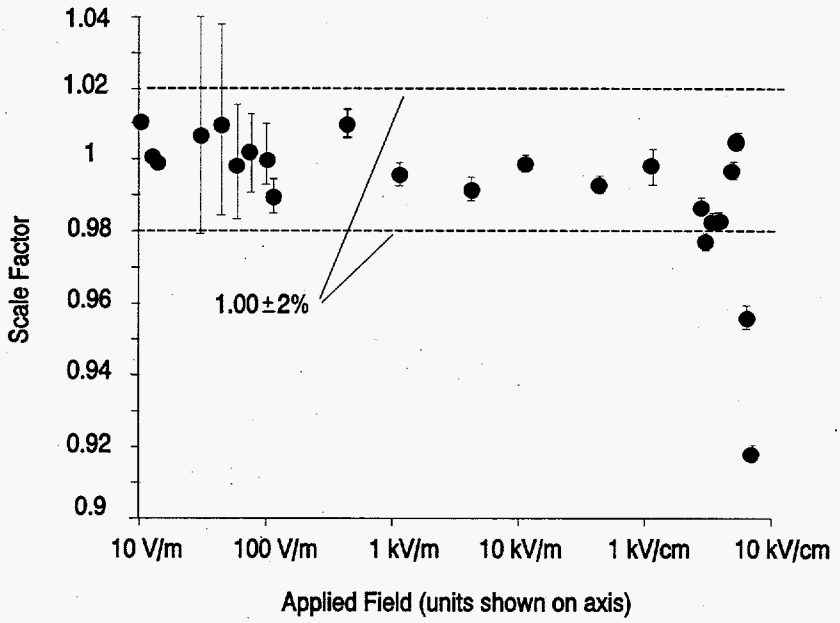


Figure 1-22. Scale factor of ac meter as a function of applied field

The vertical bars in Figure 1-22 represent the effect of changing the last digit of the display. The actual measurement uncertainty is somewhat larger.

Field values below 10 V/m can be obtained, but are of little real use. The decrease in sensitivity above about 6 kV/cm is due to saturation of the data channel,in particular the probe VCO. Between these two extremes, the linearity is excellent.

The dotted lines represent 2% of the actual value. Most of the data lie within 2% of the true value. Knowledge of the value of a field to an accuracy of better than 2% is sufficient for all practical purposes.

This field meter has three advantages over earlier field meters: it is much smaller, it can be connected to a data acquistion system and it is spherical, so it can be calibrated from first principles, without the use of a test cage.

Since it was made, the JPL meter has been used outside JPL for several applications. It was used to calibrate other field meters, and to measure the fields inside animal pens, as part of bilogical studies at Bonneville Power Administration. It was used to measure the field near insulators, as part of a research program at Hitachi in Japan (for which a 50/60 Hz switch was added), and it was used to investigate the field of electric blankets of various types, at the laboratory of the Food and Drug Administration near Washington, DC. It has also been used in a study of the shielding effectiveness of lineman's clothes as part of a student project at Cal Poly, San Luis Obispo.

It is still working, and is available for use.

1.2 MAGNETIC FIELDS

For many years, the magnitude of the electric fields near power transmission lines was of interest to researchers interested in biological effects. To gain insight into any possible interactions between power line electric fields and biological subjects, researchers want to map the electric field in air around the biological subject using probes that do not affect that field. The acquisition of data describing the way that the power line field is altered by the presence of the biological specimen, and a variety of other measurements, has led to the development of mathematical models of the specimen. Some of these models are quite refined, and it is possible, in some cases, to predict the currents induced in various parts of the body by exposure to a 60 Hz power line field (Deno, 1977; Kaune, 1986).

The shape of most field meters used in power line measurements was governed by the need to include the readout device in the probe (see, for example, Deno and Zaffanella, 1975). Accordingly, the side of the probe facing the operator was flat, and the probe overall was box-shaped. Since the DOE/JPL ac field meter had a readout at the remote end of a fiber optic cable, the shape of its probe could be chosen arbitrarily. A spherical shape has two advantages: its sensitivity can be calculated from theoretical considerations, and it results in minimum disturbance of the field. Therefore, this shape was chosen for the ac field meter probe. Details of this and other spherical probes are given by Kirkham, Johnston, Jackson and Sheu, (1987).

During 1987, a report on the biological effects of power line fields was published by the New York State Power Lines Project (Ahlbom, Albert, Fraser-Smith, Grodzinsky, Marron, Martin, Persinger, Shelanski and Wolpow, 1987). The report suggested the possibility of a link between the magnetic fields associated with the transmission, distribution or use of electric power and health effects, particularly an increased rate of certain kinds of cancer.

The linkage between the power line as cause and the increased cancer as effect was far from established and certainly not understood. Nevertheless, the report aroused considerable interest in the topic of magnetic field effects. Further investigations must be made into the possible relationship between power line magnetic fields and health effects, and certainly part of these investigations must be experimental. This provided the motivation for our involvement in the topic of magnetic field measurement.

There are several differences between the magnetic fields associated with power lines and the electric fields associated with them. First, the electric fields are relatively constant in time, because most power transmission and distribution systems are operated at a constant voltage. This contrasts with the situation of the magnetic field, which fluctuates on a moment-to-moment basis and with a daily cycle, because it is determined by the load on the power system. Similar observations apply to both transmission and distribution circuits.

The changing nature of the magnetic field means that the "exposure" of a biological subject cannot simply be determined by multiplying the exposure time by the value of the field established by a spot measurement. Some kind of integrating meter which could furnish, for example, an amplitude probability distribution (APD) of the exposure to the field would seem

to be a reasonable way to bound the problem of knowing the exposure. However, an exposure meter such as this is simply a probe (of the kind already developed at JPL for the measurement of electric fields) connected to some kind of classifying data acquisition system.

Consequently, it was thought worthwhile to adapt the ac electric field probe and measuring system to the problem of measuring magnetic fields. This proposed modification would then result in the availability for experimental investigations of an electrically-isolated magnetic field probe.

A second important difference between electric and magnetic fields is in the dynamic range of the measurement. Electric fields greater than about 1 MV/m (10 kV/cm) cannot be measured because breakdown of the air around the probe itself will be a limit. At the other extreme, electric fields smaller than about 10 volts per meter are not of interest in an outdoor environment because typical fair weather ambients are in the range of 100 V/m. Therefore, equipment with a voltage dynamic range of 100 dB can be used to cover the entire span of interest for power line electric fields. The equipment developed by JPL for the US Department of Energy has shown that it is possible to make measurements over this kind of dynamic range without making any changes in the probe itself. In our instrument, the necessary range switching is accomplished entirely within the receiver unit.

A similar dynamic range is required of the magnetic field meter for use in biological studies. The magnetic field near a household appliance, for example, can approach 1 mT (milliTesla: 1 milliTesla = 10 Gauss). The peak field under a high voltage transmission circuit could be in the order of 70 μT and the ambient 60-Hz background (in the remote countryside?) is probably about 10 nT, superimposed on the 30 μT steady-state Earth's field. To cover a useful range, a dynamic range of the same order as the electric field meter would seem to be indicated. For such measurements, a single-range probe of the kind developed for the measurement of electric fields can furnish the required dynamic range. However, magnetic fields can cover a much wider range of values. Should fields higher than a few mT be present, for example, near machinery (domestic or industrial), it would be necessary to de-sensitize the probe.

A third distinction between electric and magnetic fields is in the spectral purity of the field. The ac electric field is predominantly 60 Hz (in the US) and the distortion of that field is limited to a very few percent. Distortion of the voltage waveform corresponds to distortion of the 60 Hz source, which is normally very difficult because the power system generators are designed for constant voltage operation, and the power system has a rather low impedance. In some cases, the limit on voltage distortion is a statutory one, but in most cases, the power system is strong enough to prevent more than a few percent total harmonic distortion appearing on the voltage.

The same cannot be said for the magnetic field. The nonlinearity of the magnetization curve of iron causes the current waveform of some motors and transformers to be rich in harmonics, predominantly the third harmonic. Harmonics are also caused by switching devices, such as speed controllers for small electric motors, and light dimmers. These can cause second and higher order harmonics extending into the RF region. In terms of their energy content, the predominant distortion is the third harmonic caused by iron in transformers and motors. To be most useful, a magnetic field measuring system must be capable of furnishing some information

about the spectrum of the magnetic field, or at least about its low-order harmonics.

The dc magnetic field of the earth typically has a magnitude of 30 μT, a value comparable with that at the edge of the right-of-way of high voltage lines. Since this pre-existing field is so large, and since HVdc lines are so few in number and produce the same kind of field as the naturally occurring earth's field, there seemed little to be gained from our developing a dc version of the magnetic field meter. (In any case, there are a number of magnetometers that can make measurements of fields of this size, available for use in a variety of applications.) We therefore concentrated on ac measurements.

1.2.1 SENSOR DESIGN

1.2.1.1 PROBE

The hybrid integrated circuit for the magnetic field meter measurement probe differs from its predecessors in two ways. First, the electric field meter system required an electrometer amplifier as a first stage, with zero input impedance, whereas the magnetic field meter required an infinite input impedance if calibration from first principles was to be possible. Second, since this meter was not part of a rotating system, mechanical balance was not an important consideration in the layout of the components on the hybrid substrate.

The circuit is conventional. A pickup coil is connected to the noninverting input of the first stage, an operation amplifier. This ensures that the source will not be loaded: the input impedance of an amplifier configuration is very high. The amplifier is arranged to have an integrating characteristic to compensate for the fact that the induced voltage is proportional to the rate of change of the magnetic field. Consequently, the output of the first amplifier is a voltage replica of the field to which the coil is exposed.[1]

There is no range switch on the probe. Although simple switches were tried during the

[1] It would have been very easy to arrange for the first stage to have a flat frequency response, so that the output of the first stage would represent the voltage induced by the field to which the coil was exposed. Indeed, it seems that there was at least one magnetic field meter available on the market that had this characteristic (Leeper, 1987). However, such a characteristic would provide the overall meter with essentially a rising output with frequency, and would unjustifiably emphasize the harmonics in the field. We could see little point in choosing to emphasize these harmonics, and it was our opinion that the rising characteristic which came about because of the derivative nature of the induced voltage was an accident of physics and not to be regarded in any way as a desirable feature of an instrument which purported to measure magnetic field. Consequently, the output of the first stage is a replica of the field to which the coil is exposed, and contains fundamental and harmonics in the same proportion as the magnetic field.

development of the probe[1], the non-inverting integrator design did not lend itself to gain changes of much above 10:1, a scarcely worthwhile increase in dynamic range. The number of turns in the pickup coil, in combination with the gain of the first stage, sets the sensitivity of the instrument.

The first stage signal then modulates the voltage controlled oscillator (VCO) in a way exactly analogous with our earlier electric field meters. Short pulses are generated in the phase-locked loop chip, amplified by the light emitting diode driver stage, and coupled from the LED (which is not part of the hybrid circuit) into the optical fiber for communication to the receiver. The center frequency of the VCO is approximately the same as the VCO in the ac electric field meter, namely 10 kHz.

One improvement in the design of this hybrid was the inclusion on the hybrid substrate of the voltage regulator diodes. These serve to stabilize the power supply voltage for the circuitry on the hybrid at ± 2.5 volts. Note that no dropper resistors are included for these diodes. The photodiode arrays driving these circuits have sufficient internal resistance, and the current handling capability of the regulator diodes is not exceeded[2].

A feature of this hybrid is that the header to which the ceramic substrate is attached is not made of the conventional material. That material is magnetic: usually the header is made of steel and the pinouts are Kovar, another magnetic material. Since the presence of material with a high permeability would certainly distort the field that we were trying to measure, special nonmagnetic headers (made of fiberglass) and pinouts were fabricated for this application.

Mechanical Construction

The construction of the magnetic meter probe was based on the design of the electric field meter built earlier. The probe was spherical, with the hybrid IC about at the center of the sphere. Two phenolic formers were used to hold the IC in place. To them were attached the "paddle" that held the probe assembly on the end of the support pole, and the two hemispherical electrodes. In the magnetic meter, the electrodes were not part of the sensing system, they merely control the electric field intensification. The sensing element is a coil of fine wire wound in a groove in one of the phenolic formers.

---

[1] One such switch was a very simple mechanical device. A screw was inserted through the shell of the probe and grounded the resistor which controls the gain of the first stage. This causes the gain of the stage to increase. However, when this approach was used to change the gain of the stage by 100:1, the frequency response of the integrator was far from ideal at the low gain setting. A revised probe design, using an FET follower input stage and an inverting-mode integrator, could be combined with such a range switch to provide an additional 40- or even 60-dB dynamic range.

[2] During 1987, we found a range of commercially available diode arrays. These were made in silicon, with dielectric isolation. Continued problems with producing our own GaAs photodiode arrays made use of these components the obvious choice. Because their design has been optimized, the efficiency is comparable to our earlier gallium arsenide arrays. The manufacturer was DIONICS Inc, of Westbury NY.

A photodiode array, made by Dionics Inc, was mounted on the paddle, at the rear of the hybrid IC, and the fiber feeding energy to it was fixed a few mm away. The paddle arrangement was rather simpler than in the case of the earlier electric field probe because there was only one photodiode array, furnishing power to both the positive and negative supplies. The LED which sent the output data to the receiver was mounted on the opposite side of the paddle. Figure 1-23 is a cross-section showing the construction of the probe.

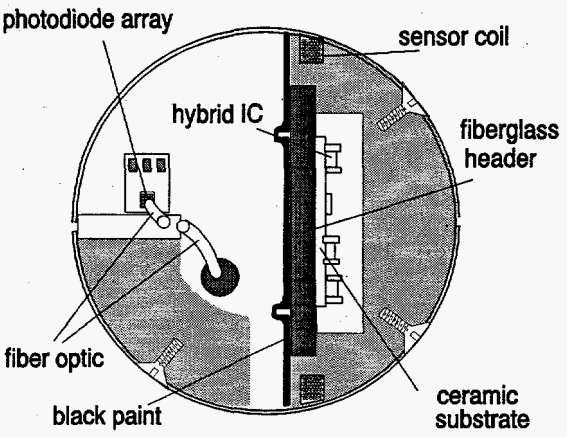


Figure 1-23. Cross-section through magnetic field meter probe

1.2.1.2 RECEIVER

In the earlier electric field meter, the reconstructed signal was filtered by a state variable filter with a two-pole low pass characteristic and a cut-off frequency of 500 Hz, before being passed to a pair of gain-adjusting potentiometers. This filter is replaced in the magnetic meter by the multi-pole filter used to measure wideband (up to 1 kHz) energy. The gain adjusting potentiometers allow the subsequent amplifier to have the appropriate gains for both electric and magnetic field probes over its four switch positions, each of which represents an order of magnitude difference in gain.

It was mentioned earlier that the synchronous detection which had been used to provide the wide dynamic range for the electric field measurements would probably be unsatisfactory in the case of the magnetic measurements, which were expected to be highly distorted. The phase-locked loop system used in the synchronous detector would be able to lock onto the distorted waveforms. However, the phase relationship between the fundamental and the harmonics could not be specified in advance, and hence the average indicated by the output of the synchronous detector would not, under most circumstances, adequately reflect the magnitude of the fundamental of the field being measured, let alone its harmonics.

Instead, a filter arrangement was designed to permit separation of the fundamental, second and third harmonics, and their separate measurement. The broadband energy is also measured. A block diagram of the circuit board which accomplished this is shown in Figure 1-24.

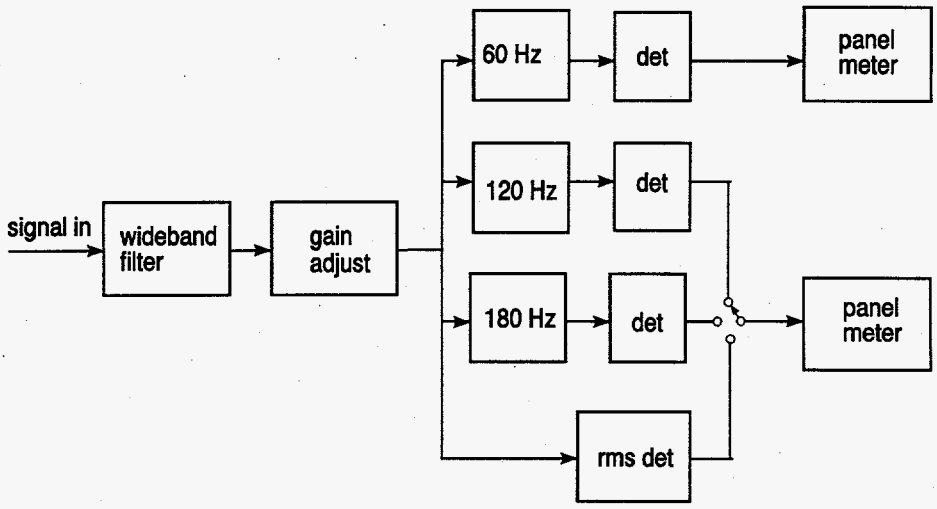


Figure 1-24. Block diagram of magnetic field meter

The bandpass filters shown in Figure 1-24 are centered at 60, 120 and 180 Hz. Each was constructed as a six pole Butterworth-characteristic filter. Each pair of poles in the filter was realized by means of a state-variable implementation.

The broadband filter, designed to be flat between about 10 Hz and 1 kHz, was realized using a slightly different configuration. The low frequency cut-off of the filter was accomplished by a capacitor-coupled amplifier, providing a single-pole response. This was followed by three state-variable filters, configured in the low-pass mode. The high frequency roll-off created by the six poles of the filter was thus quite sharp. This helped to minimize the noise contribution of the optical fm carrier.

The broadband filter was placed ahead of the gain-adjusting stage in order that the dynamic range of the amplifier not be exceeded by the noise spikes in the signal. This would lead to nonlinear operation, and the device would be impossible to calibrate. As is good practice to prevent ringing in high Q stages, the three state variable filter sections which were cascaded to construct the six-pole Butterworth filters each began with the lowest-Q pair of poles.

1.2.2 CALIBRATION

1.2.2.1 CALIBRATION FROM FIRST PRINCIPLES

Since the input impedance of the first stage of the probe is extremely high compared to the source impedance of the coil, the calibration from first principles can be done quite accurately. One need only know the area (or radius) of the coil of wire, the number of turns N, and the frequency (which for calibration purposes can be taken as 60 Hz) and then apply the relationship:

$$e = N \frac{d\Phi}{dt} \tag{1.11}$$

where e is the voltage induced by a changing flux ϕ. The flux may be assumed to be sinusoidal with time for calibration purposes, and the total flux ϕ is given by the flux density B, for which the instrument is being calibrated, and the area of the coil. If the rms value of the flux density is used, Equation 1.11 will yield the rms value of the induced voltage. Thus, with 60-Hz excitation,

$$e = 1184.35\, B\, r^2 \tag{1.12}$$

for each turn of the coil. Here e is in volts if B is in Tesla (or Webers/m^2) and r is in meters. A known voltage e is applied to the probe and the gain of the receiver is adjusted until the meter reads the appropriate value B of flux density.

In our design, we have $r = 8.5$ mm, and a coil of 100 turns, so that

$$e = 8.557$$

for a field of 1 Tesla. Since we are typically dealing with fields smaller than this by a factor of 10^6, the input voltage to the hybrid IC is typically in the order of μV.

The calibration procedure is as follows. To represent a field of 1 mT, which is close to the largest field that the meter can read, a voltage of 8.56 mV, at 60 Hz, is applied to the input of the probe, in place of the pickup coil. The gain switch is set to the least sensitive position. The overall-gain adjusting potentiometer is set so that the outputs of the 60-Hz display and the rms detector indicate 1 mT. The frequency is then adjusted to 120 and 180 Hz in turn, and the corresponding gains adjusted so that the readings are one-half and one-third of the 60-Hz value respectively.

Internal gains are designed so that the readings on the displays are as large as possible, to maximize the usefulness of the display. The range switch has four positions, each one order of magnitude more sensitive than the last. In the least sensitive (highest field) position, a full scale reading corresponds to 1.999 mT. In the most sensitive (smallest field) position, a full scale reading corresponds to 1.999 μT. It is possible to measure fields in the order of 0.1 μT.

1.2.2.2 EXPERIMENTAL CALIBRATION

As in the case of the electric measurement, it is good practice to verify the calibration experimentally. Just as a uniform electric field can easily be generated by applying a known potential to two parallel plates, the Helmholtz coil arrangement can be used to generate a magnetic field which is known and uniform over a small region, although the region of uniformity in the magnetic case is not as large as in an electric field cage of comparable volume.

Our Helmholtz apparatus was about 0.3 meters in radius and length. Consequently, the field at the center of the apparatus is uniform over a region about 5 cm in extent, which is certainly large enough to permit calibration of the small probe used in our field meter system.

Uncertainties in Calibration

It can be seen from Equations 2-1 and 2-2 that uncertainties in the area, frequency, and number of turns of the probe coil are simply additive in their effect on the calibration accuracy. With care, the number of turns can, of course, be accurately determined. The frequency is known to be 60 Hz with great precision, so that the predominant source of uncertainty becomes the accuracy with which the area of the coil can be known. Since very fine wires (50 AWG) are used in its construction, the coil has only a small, although not infinitely small, extent in the radial direction. The wire used has a diameter of 0.0254 mm (0.001 in.), so that if the entire coil is constructed in a neatly wound arrangement of, say, two layers, the coil extends for about 1% of the radius. Since the coil thickness is such a large fraction of the radius, the combined uncertainty of measurement of the radius and assuming that the coil is located entirely at its center can be appreciable, say 1%.

Uncertainty in the calibrating voltage must be added to the total uncertainty, and may be on the order of 0.5%. If we consider all the sources of uncertainty in the calibration, it is unlikely that a first-principles calibration uncertainty better than 2% is possible.

In the case of the Helmholtz coil calibration setup, we obtain the following expression for the field at the center of the coil pair:

$$B = 4\mu_0 \pi r^2 Ni \left(\frac{1}{(r^2 + (r^2/4)^{3/2})} \right) = \frac{32\mu_0 \pi Ni}{(5)^{3/2} r} \tag{1.12}$$

where μ_0 is the permeability of air, and N the number of turns in each coil of the Helmholtz apparatus. Hence the uncertainty of the field at the center depends upon the number of turns, on the current, and on the radius, a parameter which was obtained only with moderate precision in our apparatus. (The method of construction that was used was such that the coils are only approximately circular.) Thus, while it is possible to know the number of turns precisely, the current can probably not be measured with an uncertainty less than one-half a percent in our laboratory, and somewhat larger errors are certainly accumulated in the dimensions of the device.

It may be concluded, then, that the Helmholtz arrangement provides a convenient way to check the accuracy of the first principles probe calibration, but the sources of uncertainty in the Helmholtz coil calibration are of similar magnitude to those in the first-principles calibration of the probe.

Light Sensitivity of Hybrid

An unexpected result was obtained when the first nonmagnetic hybrid IC was energized by means of the laser. Compared to the power consumption of the IC, a considerable amount of optical energy is inserted into the probe. While some of it is absorbed by the diode array, some of it is reflected and illuminates the interior of the hemispheres. It was discovered that the hybrid integrated circuit, now mounted on a fiberglass header, was light-sensitive.

Apparently enough of the infrared energy from the laser was penetrating the fiberglass to appear on the circuit side of the substrate. The resulting photons acted as input current in the sensitive input stages of the integrating operational amplifier. Under some conditions, it was found that enough energy was available to cause the integrator stage to saturate. At this point of course, the entire probe ceased to function. The solution was to paint the exterior of the fiberglass header with an opaque paint.

Dynamic Range

In contrast with our earlier electric field meters, the magnetic field meter has a wide frequency response and does not make use of synchronous detection. These two features make the question of providing the instrument with a wide dynamic range particularly challenging. A wider bandwidth can ordinarily be expected to provide a greater amount of noise in a system, and the use of synchronous detection enables the recovery of signals from under a large noise. In the present design, the dynamic range can be expected to be smaller than that of the electric field meters, particularly as far as the wideband measurement of the signal from the magnetic field probe is concerned.

Probe Saturation

The maximum field that can be measured is limited by saturation in the system. The system was designed so that the probe saturated before the receiver, so that for ordinary conditions the amount of modulation on the optical link was maximized. This maximizes the dynamic range, if noise is introduced in the link.

The design goal for our field measuring system was a maximum field of 2 mT. (Actually, the displays would go blank for a reading greater than 1.999 mT.) To ensure linearity up to this point, the probe was designed to saturate at about 2.3 mT. Because the first stage of the probe is operated as an integrator, its output is larger at lower frequencies. Therefore a 60-Hz signal will cause saturation first. The trade-off here is between the number of turns in the coil, and the gain of the first stage amplifier.

The use of a noninverting integrator led to a situation in which the response departed from that of an ideal integrator at high frequencies. The frequency at which this departure occurred was deliberately chosen to compensate for a corresponding high-frequency roll-off in the optical link. This decision fixed the gain of the integrator, with the result that the only free parameter to establish the proper saturation field was the number of turns in the coil. In our particular case, a 100-turn coil producing about 17 mV at 2 mT was used. A signal of about 20 mV at 60 Hz

was sufficient to cause clipping in the integrator output.

Noise Floor

At the other extreme, the noise level in the system sets the lowest signal that can be measured. A wideband signal ordinarily contains more noise than a narrowband signal. Because of this, the wideband measurement of the magnetic field can be expected to reach its noise floor ahead of the single-frequency measurements at 60, 120 and 180 Hz. In fact, if the noise is Gaussian, the difference in dynamic range between the various measurements can be obtained simply by examining the difference in bandwidths. For example, two filters whose bandwidths are B_1 and B_2 can be expected to have dynamic ranges differing by a factor of B_1/B_2, assuming that all of the noise is entering at the input to both filters.

The 60-Hz filter is about 1 Hz wide. The wideband and narrowband filters thus have bandwidths in the ratio of approximately 1000 to 1. It was expected that the narrowband measurements would have a dynamic range two orders of magnitude (40 dB) greater than the wideband measurement. In fact, the noise floor of the narrowband measurements is better than that of the wideband, but by a rather smaller amount than expected.

In our field meter, the noise floors of the various filters differed by a factor that seemed closer to $(B_1/B_2)^{1/2}$ than it did to B_1/B_2. Therefore, to maximize the improvement in dynamic range due to the narrowness of the single-frequency filters, they must be as narrow as possible, within the limitations of realizable audio filters. The use of even higher-Q filters is indicated, but these may be difficult to implement.

1.2.2.3 EXPERIMENTAL RESULTS

The dynamic ranges of the wideband measurement and the 60-Hz measurement are shown in Figure 1-25, which compares the experimentally determined error in each measurement as a function of applied field. By using an external voltmeter, it was possible to show the instrument output above the 2-mT at which the panel meters indicate over-range.

The graph shows an apparently constant error of less than 1% for fields between 2×10^{-6} T and 2×10^{-3} T. This approach was used to represent the uncertainties caused by the range resistors and the resolution of the display. There is not a constant error, rather the line indicates that 1% is reasonable accuracy to expect from the measurement. Above 2 mT, the error increases very rapidly. This is the effect of clipping in the probe. The error is such that the meter reads consistently low. Below 10^{-6} T for the 60-Hz measurement, or 10^{-5} T for the broadband measurement, the error increases gradually, as noise makes an increasingly large contribution to the reading. The meter reads high.

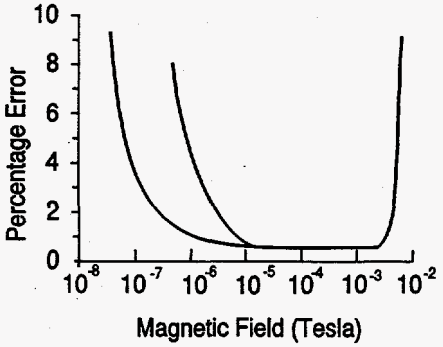


Figure 1-25. Dynamic range of field meter

The dynamic range of the wideband measurement is about 3 orders of magnitude (actually 66 dB), from 1 μT to 2 mT (or 10 mG to 20 G). The dynamic range of the 60-Hz measurement is somewhat over 4 orders of magnitude (about 88 dB), from about 80 nT to 2 mT (or 0.8 mG to 20 G). These results show that, despite the narrowness of the 60-Hz filter, the measurement is only a little over 20 dB more sensitive than the wideband 1-kHz measurement. If a more sensitive instrument is needed, the 100-turn coil could be replaced by one with more turns. The scaling is simply linear in the number of turns.

The source of the noise in the measurements was investigated experimentally. A high-quality signal source was substituted in turn for the digital signals at various points in the system, with no modulation. It was found that the system noise was essentially unchanged when the probe signal was replaced, or when the PLL demodulator was driven directly by the signal generator. The probe and the link were therefore eliminated as noise sources. When the PLL output was replaced by a square wave pulse train, the system noise was considerably reduced. It was concluded that the predominant source of noise in the system was, as may be expected, phase jitter in the PLL demodulator. There was very little that could be done to reduce this effect. The next largest source of noise seemed to be shot noise in the amplifiers of the flat filter. Since all subsequent stages see this noise, it is important that it be minimized. High quality amplifiers were used in our prototype.

Finally, it seems that the noise is not Gaussian. Its spectral density seems to correspond more to pink noise than white; equal energy is contained in any two octaves of the spectrum, rather than in any two regions of equal bandwidth. The issue was not pursued further: a complete investigation of the noise properties of the PLL was beyond the scope of the present study.

Calibration Accuracy

The dynamic range tests shown above were performed "synthetically." As in the calibration procedure, a signal generator was used to inject a voltage into the probe. This was the only practical way to obtain readings near the low end of the dynamic range, since the magnetic field noise floor of our laboratory was higher than the electronic noise floor of the equipment. Further, the output was monitored on a separate voltmeter, so that readings above the clipping point of the probe (and the blanking point of the displays) could be obtained. To provide

verification of the accuracy of the calibration, the probe was inserted into the Helmholtz calibration apparatus.

Fields in the order of 10^{-4} T were used as spot checks on the calibration. Such fields were at least two orders of magnitude greater than the magnetic field noise floor of our laboratory. No difference greater than 1% was found between the calculated field and the measured field. It was concluded that the synthetic calibration we used was at least as accurate as an experimental one would have been, and had a much wider dynamic range.

Example of Field Measurement

For demonstration purposes measurements were made, using this system, of the magnetic field near a small 1500-W hair dryer. In the past we have presented field measurements of this kind in the form of lines representing the vectors of the field. On this occasion it was thought to be more useful to show the field as contours of equal magnitude. In order to develop such a contour map from the measured data a computer program was written. Representative outputs of this program, showing the contours of the magnetic field at the fundamental frequency, and the third harmonics are shown in Figures 1-26 and 1-27. The program is described by Kirkham and Johnston (1988).

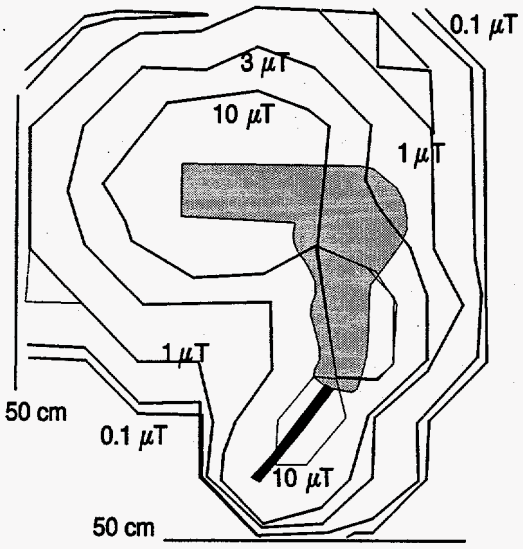


Figure 1-26. 60-Hz magnetic field around hair dryer

The fundamental field, shown in Figure 1-26, has the same general shape as the hair dryer. There seem to be contributions to the field from both the motor and the heater coils, as may be expected. The flexible wire by which the hair dryer is energized has only a relatively small external field, presumably because the two conductors are close together.

The data shown in the Figure were measured in the plane of the device. The contour that appears to cross the hair dryer results from the plotting program's interpolation across that region.

The outer field contour, corresponding to 0.1 μT, is shown very close to the contour for 0.3 μT.

In fact, while it is a small field, this contour is above the noise floor of the field meter, and evidently represents the effect of the ambient magnetic field in our laboratory.

It may be estimated from the figure that the peak field at 60 Hz was in the order of 20 μT (200 mGauss). In fact, values of this order were measured, but do not appear in Figure 1-26 because of our choice of contours.

There was a small 180-Hz component to the field. This is shown in Figure 1-27.

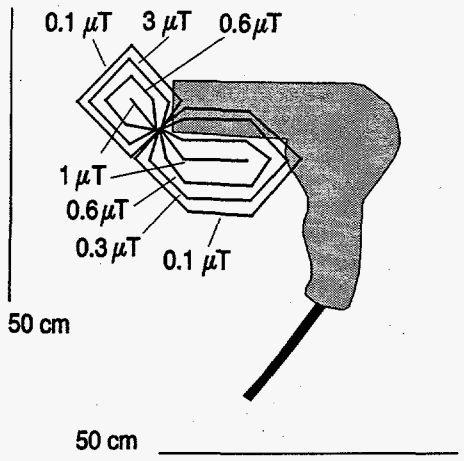


Figure 1-27. 180-Hz magnetic field around hair dryer

It was thought that it might be interesting to measure the magnetic field in an office. Most convenient was the office of the author of this report. This office contains a personal computer and a laser printer, on which the original of this document was prepared. A 60-Hz field of about 0.1 μT (1 mG) was measured in the area in front of the terminal. However, investigation showed that this field was originating in the printer, located nearby. A 60-Hz field of about 1 μT was measured near the printer, and values as high as 7 μT were observed at the surface of the printer itself. No measurable fields at 120 or 180 Hz were observed.

Perhaps it should be stressed that we attach no particular significance to the results presented here. These experimental data are given merely to indicate the type of measurement that can be made with the prototype field meter. While none of the measurements actually required the use of an electrically isolated probe, the results do show that fields that are quite small in magnitude and that are not spatially uniform can be measured with the device.

It is concluded that the small optically-powered probe and its filtered measurement receiver provides a system for the measurement of magnetic fields over a practical range of values, with the advantage of electrical isolation.

Section 2: OPTICALLY POWERED ELECTRONICS

SUMMARY

In many respects, the communication needs of the power distribution system are unique. Communications must reach all parts of a system whose configuration is likely to be a complex mesh, and must do so even if part of the network is damaged. A master-slave approach to monitoring and control is preferred, with a polling cycle taking just a second or two to get measured data from, say, 100 locations. A set of protocols was designed by the Communication and Control Project at the Jet Propulsion Laboratory specifically to address the needs of power distribution.

The protocols, called AbNET, can be widely used for data acquisition. This part of the Final Report discusses the background against which AbNET evolved, and details the combination of unusual solutions AbNET uses.

In communication terms, AbNET is a connectionless distributed network, combining flooding and antibodies to maximize fault tolerance and control link traffic, polling to control medium access and zones to permit communications to correspond to geography. The combination is useful because it provides maximum fault tolerance while a the same time meeting the requirements of the distribution system.

SUMMARY

With the development of efficient optical to electrical converters, optically transmitted power could be used to energize electronics for power system measurements. A system powered by optical energy would avoid the need to replenish batteries, or to extract power from power lines.

Work done at JPL has included the development of photovoltaic arrays in both silicon and gallium arsenide, and the development of very low power electronics (both analog and digital) capable of operating with the small amount of power available.

This work is descibed in this section of the Final Report.

Fiber optics can be used to supply power for electronic sensing, producing a sensor system that has the isolation benefits of fibers and the flexibility and general applicability of electronic sensing. Although solid-state lasers can deliver considerable power, and GaAs photocells can convert much of it into useful volts and amps, a cost effective system requires that the power demand of the sensor system be minimized. Some of the trade-offs involved are discussed. Experimental results showing optically powered instrumentation (strain gauge, accelerometer) are presented.

2 OPTICALLY POWERED ELECTRONICS

2.1 INTRODUCTION

The use of optical power for electronics was about a decade old (DeLoach, Miller and Kaufman, 1978; DeLoach and Gordon, 1981; Ohte, Akiyama and Ohno, 1984; McGlade and Jones, 1984; Hall, 1986; Schweizer, Neveux and Ostrowsky, 1987) before a commercially successful application of the technology was found (Adolfsson *et al*, 1989) in measuring current in the high voltage part of a power system. In this application, optical powering allows the instrumentation designer to create a sensor that has the advantages of conventional electronics, while having the external attributes of being optical. The only connections to the measuring device are optical fibers, that can easily be contained inside a composite insulator, so that the sensor can be operated at line potential.

The system cost is thus not dominated by the cost of insulation, as it often is in conventional current measurements in high voltage systems. Indeed, it seems as if a good deal of energy was put into researching this topic because of the apparently simple economics. One might speculate on the argument that a research group used to obtain company IR&D funding. A fiber-based current measurement system could be manufactured for so much less than a conventional CT that it would surely be a profitable product. See Figure 2-1 for example.

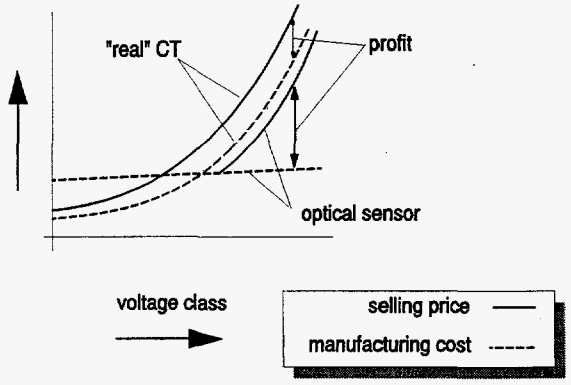


Figure 2-1. Conventional CT fiber-based CT costs compared

Because the cost of insulation rises rapidly with system voltage, fiber-based technolgy becomes less costly at some voltage level[1]. Above that point, the selling price of a fiber-based system

[1]One of the reasons that the cost of conventional insulation rises so rapidly is that insulation has to be provided for what is caled the Basic Inpulse Insulation Level (BIL) of the system, and this is much higher than the system voltage. While the ratio of BIL:system voltage is lower at high voltages, it is still sufficiently high that a transmission level CT would, for example, have to be designed for a BIL over a million volts.

can be increased to improve margins, without danger that a conventional device could undercut the price. The reader will notice that Figure 2-1 has no labels on the axes. This is deliberate: the crossover point depends in practice on so many assumptions that it would be misleading here to identify a particular cost or a particular voltage level.

The persuasiveness of this argument has resulted in the development of many optical current measurement systems. A number of quite different optical current transducer (OCT) systems are now on the market or undergoing field trials. They all use optics to isolate a high-voltage part of the system from a grounded part, as illustrated in Figure 2-2.

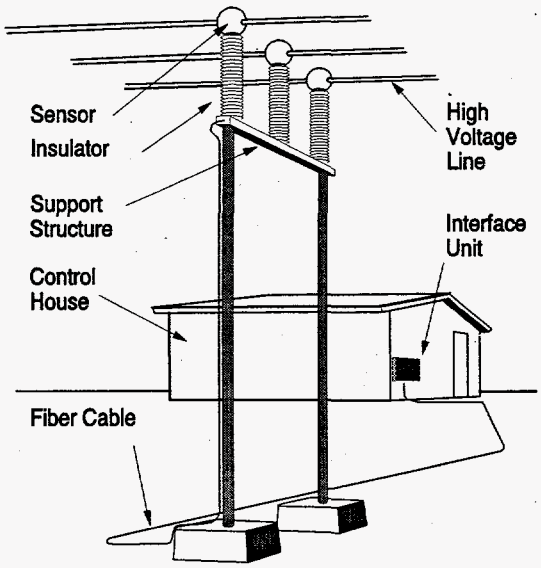


Figure 2-2. Essential elements of OCT system

It would be misleading to think of an OCT as an optical current *transformer*. OCTs are optical and electronic measurement systems. There are actually many different ways to make an OCT, and most are not based on the transformer principle.

Because an OCT is optical and electronic, a fundamental way an OCT differs from a CT is in the signal power involved. In all the OCTs considered here, the current being measured is represented, as it is transmitted from high potential to ground, as modulated light. In a CT, the secondary signal has a power level of several watts. The power in the optical part of the OCT is typically a few μW.

Diversity is present in all elements of the system shown in Figure 1. The sensor itself may be optical or electronic, so the high-voltage part of the system may be active or passive. The insulator may be ceramic or polymer—it may be used to support the OCT or it may be suspended from it. Typically, but not universally, it contains an optical fiber carrying the OCT output signal. The way the current information is encoded in the fiber varies from system to system. This affects the design of the receiver at ground potential. An interface unit connects the OCT system to the user device, which may be a relay or a meter or other equipment.

All optical current transducers share several advantages over conventional devices. One is that

they are light weight. The optical sensor may be much lighter than a conventional oil-filled CT of similar ratings. This lighter weight allows for savings during installation; the support structure is smaller, a smaller crane can be used, and installation time is shorter than for a conventional CT. Other advantages include noise immunity and safety: conventional CTs are known to fail catastrophically sometimes. Since the insulating part of an OCT consists only of an optical fiber and a fairly standard insulator, such a failure mode is less likely.

The optical measurement of current in a high-voltage power system has been commercialized in some applications. The performance of devices which have emerged from the laboratory in the last few years compare well with the conventional CTs that have been developed over about the last 80 years.

The most successful (in numerical terms) of the commercially available CTs has been an almost conventional CT with an optical output.

Add to a conventional CT an insulated optical information channel to replace the copper wire output, as shown in Figure 2-3, and the advantages of both conventional (well understood and characterized) and optical devices (see above) are retained.

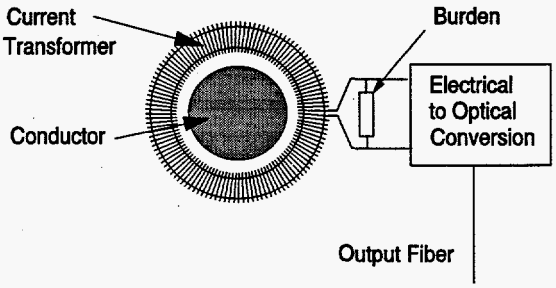


Figure 2-3. Conventional CT with optical readout

There are several methods to convert the current transformer output into optical form. In terms of systems that have got as far as field trials, there are three: one uses electronics to produce a digital optical output, one uses a Faraday sensor and one uses a piezoelectric device to convert the electrical signal to optical form.

The electronic version, produced originally by ASEA in Sweden (now ABB), was included in a series-capacitor protection package, where the entire system was at line potential. Each protection unit included many OCTs.

The general concept is shown in Figure 2-4.

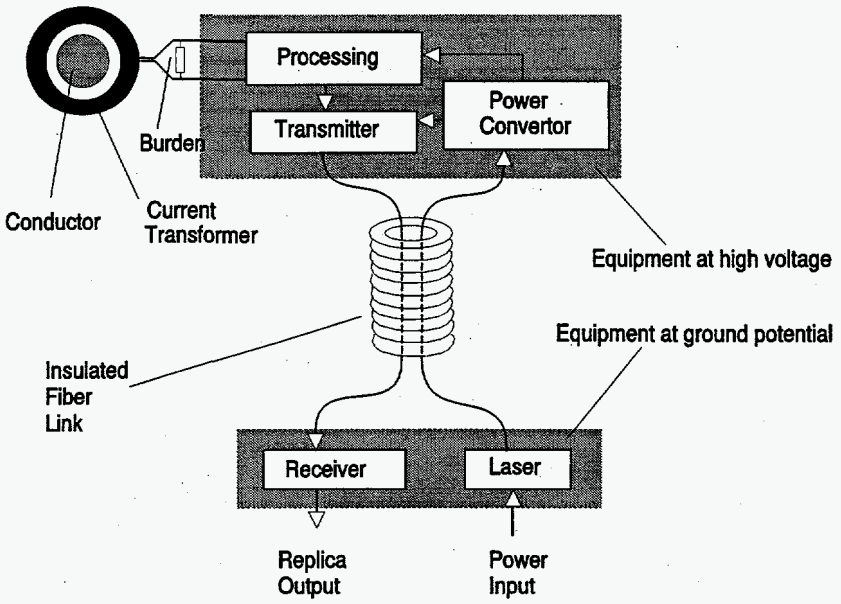


Figure 2-4. The major components of a system to measure the current in a high-voltage transmission line using a conventional current transformer and an optically-powered telemetry system

The developers at ASEA made a considerable effort to reduce the power consumption of the optically powered portion of their system, while at the same time achieving maximum performance. A new application-specific integrated circuit (ASIC) was made that included a prescaler for the A/D converter, so that the equivalent of about a 17-bit conversion was obtained.

At the same time that this system was being developed in Sweden, the JPL Communication and Control project was working on optically powered system in the US. A big difference was that our schemes were intended to be more general purpose. In the last few years, the state of the art of electronics has advanced to the point that a quite general purpose measurement system can be made, at modest cost, using off-the-shelf components. The key to the improvement is the availability of low-power microprocessors, and low-power A/D converters. Rather than using ASICs, the system that has evolved at JPL uses general purpose micrpprocessors, and can be configured to do a number of different tasks.

TRADE-OFFS

The cost of an optically powered measurement system is very dependent on the amount of power required by the remote sensor part of the system. The factors that have to be considered in the power budget are shown in Figure 2-5.

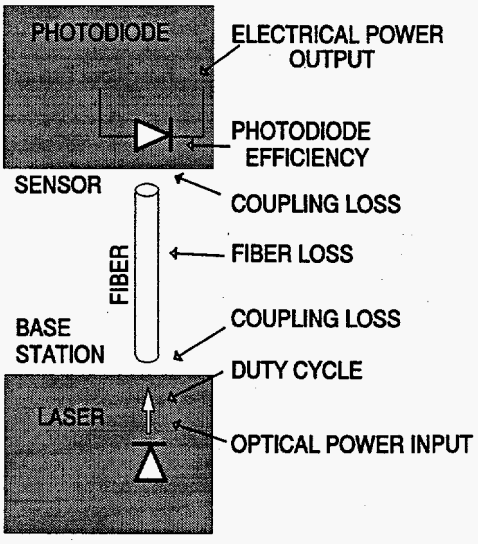


Figure 2-5. The factors that reduce the power available for the electronics

The power P_{avail} available for use by the electronics can be expressed in the following form:

$$P_{avail} = P_{laser} \times \eta_{pd} \times (1 - loss_{f\text{-}pd}) \times (1 - loss_{f}) \times (1 - loss_{1\text{-}f})$$

where P_{laser} is the laser power
η_{pd} is the efficiency of the photodiode
and the loss factors are as shown in Figure 2.

It is assumed that the laser duty cycle is 100%, and that there is no limit to the power that can be supplied to the laser.

The fiber loss is fixed by the kind of fiber and its length, and in while the loss may be large (3 dB represents 50% loss), it is probably not amenable to reduction. The same can be said for the coupling loss in and out of the fiber. The power available is thus essentially controlled by just the input (laser) power and the photodiode efficiency.

A GaAs photodiode is the best match to most lasers, and efficiencies in the order of 40% are readily achievable. A single cell produces about 0.8 V, and arrays that produce higher voltages are commercially available built into fiber connectors. However, for a low-cost system, Si diode arrays are available that can produce 10–12 volts at efficiencies of perhaps 6–8%, but at about 100 times lower cost. The Si devices we have seen have no fiber connectors, and means must be furnished to mount the array on the end of the fiber. Nevertheless, if the lower transmission efficiency does not result in a much greater laser cost, this is the preferred solution.

The range of power available is quite large. A solid-state laser can readily supply 1 W (electrical) if an efficient photodiode is used. At the other extreme, use of a low-cost CD-type laser will result in a power level of only about 2 mW after all loss sources are accounted for, less if a Si photodiode is used.

It is our objective to show that a useful system can be built to take advantage of even this low level of power. Careful design can produce a variety of practical measurement systems, based on the general concept shown in Fig 1. The overarching design rule is "If it consumes a lot of power, don't turn it on for very long." The key to achieving this aim is the use of a low-power microprocessor.

SYSTEM DESIGN

Some years ago, because of growing interest in the effects of electric and magnetic fields, we built a series of small meters to make isolated measurements of power line fields. Two of these instruments used optical power to supply the electronics because we were able to obtain photocells and photodiode arrays that were smaller than the smallest battery we could find.[8,9] This enabled us to build an electric field measuring system inside a 2-cm sphere, for example. This size was achieved by using hybrid circuit technology, and by keeping the number of parts used to a minimum. The current produced in the input of an operational amplifier was used to modulate a VCO, and the system output was FM. A dynamic range in excess of 60 dB was achieved by the use of phaselock techniques in the receiver.

More recent work would make use of digital technology. We have discussed a digital network that uses optical power.[10,11] In this paper we will concentrate on the measurement aspects of optical power. A general block diagram is shown in Figure 3.

The biggest difference between the approach shown here and our previously reported sensing work is the use of the microprocessor. The advantages include:

- It is more flexible, since the same basic design can be used to measure a large number of parameters. The microprocessor can be programmed to perform range-changing, correct for sensor nonlinearities or temperature effects, and to communicate data in engineering units back to the base station where data acquisition takes place.

- The link can be bi-directional. The remote node can be individually addressable, so it can receive data from the base station, and act upon it. This allows the possibility of synchronized sampling at different nodes, and of control at the remote location.

- It is inherently compatible with a digital data acquisition system at the base station.

- Development costs are minimized by the use of familiar electronic digital technology

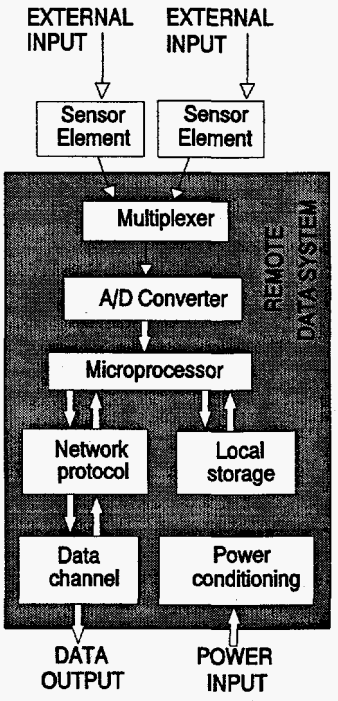


Figure 2-6. The general-purpose data acquisition node shown can be operated by optical power

Microprocessors with integral A/D converters are convenient to use, and available as micropower devices. The power consumption varies with the clock speed, which can be as low as 32 kHz. To economize on power, the clock speed can be changed dynamically, and the microprocessor can be put into a "sleep" mode. With the microprocessors used in our prototypes, a little over a thousand clock cycles are required to bring a processor out of the sleep mode, so this approach is only practicable for a system with a fast clock.

For a measurement such as strain, it is clear that the strain gauge itself would normally consume much more power than the rest of the electronics. A strain gauge with a resistance of 120 Ω supplied with 5 V consumes 200 mW, a power level that would require both a large laser and an efficient photovoltaic cell. Some improvement results from the use of a higher resistance strain gauge. Changing to 1500 Ω (a reasonable upper limit for a readily available gauge) decreases the power by an order of magnitude, but this is still too much power for a low-cost system. The solution is obviously to turn off the power to the strain gauge when it is not being sampled. The power to the analog amplifiers can be turned on and off at about the same time, under the control of the microprocessor. The exact timing depends on the clock speed, since the various analog and digital electronics must be allowed time to stabilize. Two examples are shown below. They may be regarded as presenting the typical sequence of events.

Clock speed may limit the amount of data that can be transmitted to the base station. In the first example in Figure 2-7, the results of the measurement were presented on a liquid crystal display, which was updated only every few hundred ms. In the other example, serial data were returned via an LED. Though this consumes a considerable amount of power, it can be arranged to do so for a very short time, so the average power level is held low.

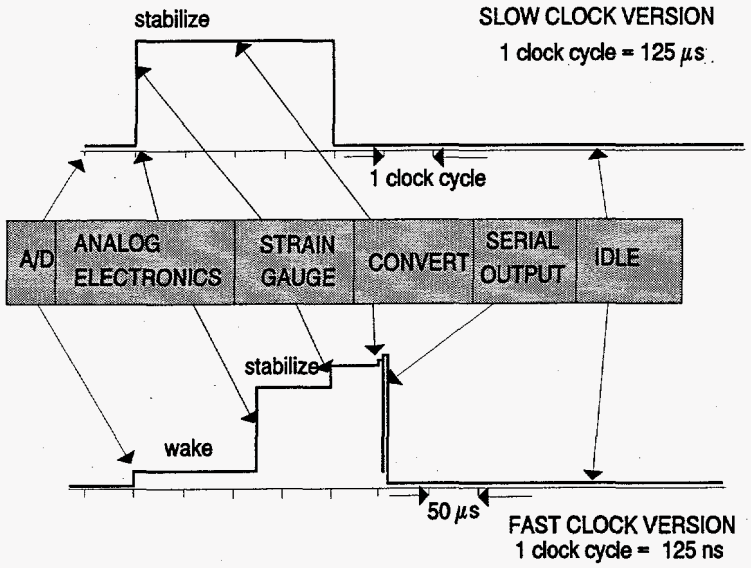


Figure 2-7. To minimize power consumption, the analog electronics can be turned off when not in use. Time must be allowed for stabilization after switch-on. Exactly how long depends on the amplifiers being used, and other factors. The time needed to stabilize is roughly constant, independent of the clock speed. Conversion time is constant at a couple of clock cycles.

RESULTS

We mentioned earlier that one of the application areas for optically powered sensing was the electric power industry, where in a high voltage system the approach offered the potential for reduced insulation cost. The parameters of interest depend on just what is being monitored: in addition to the current in a line (conventionally obtained by means of a device called a current transformer), there are situations in which it is of interest to monitor the temperature of a line, and any vibrations of the conductor. The conservative nature of the power industry makes widespread application of a novel technique (such as optical measurement) to a fundamentally important parameter (such as current) improbable[12]. (The optically powered system of ref. 7 is part of a complete package, so that the customer is not involved with the way the current measurement is done.) For our trial system we opted for a measurement that is less crucial to system operation, and that provides a challenge to the idea of optical power for electronics: the strain gauge.

Figure 2-8 shows the response of an optically-powered strain gauge system. In this example, the monitoring system comprised a 120 Ω strain gauge, an amplifier, and a microprocessor data acquisition system. To minimize power consumption, and hence the size and cost of the laser power source, the strain gauge and the amplifiers were switched on prior to each sample, and off immediately afterwards, under the control of the microprocessor. The sampling rate was 300 samples/s.

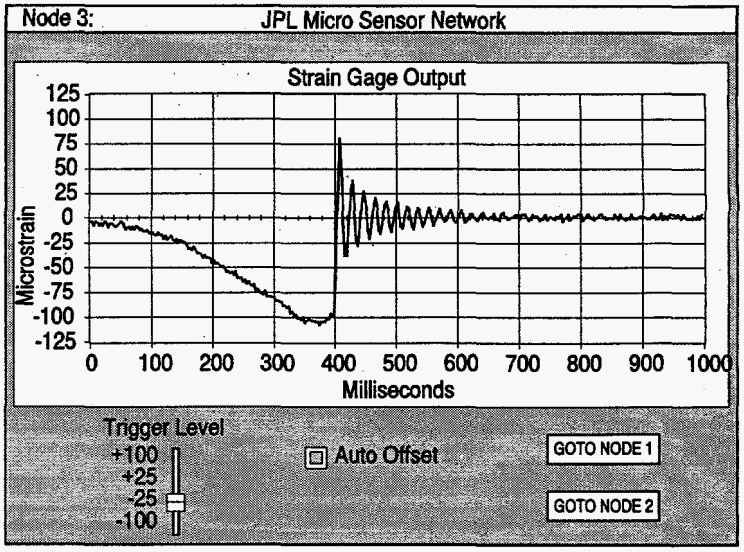


Figure 2-8. The image at right was obtained from a Lab-Windows GUI attached to the base station. The data represent the vibration of a small beam to which has been bonded a conventional resistance-wire strain gauge.

The conventional strain gauge is a very inexpensive component, that nevertheless can be tailored for quite exacting requirements (for example, matching the temperature coefficient of expansion of the material to which it is bonded so that temperature does not show as a strain). The measurement of strain is often quite challenging, however, because the signal that must be measured is very small. A reasonable sensitivity for a gauge to be used in structural monitoring is a few microstrain. This corresponds to the measurement of resistance to within a few ppm. The measurement is made possible by the use of a bridge circuit, so that, in effect, the relative values of resistors are being compared to this precision, rather than their absolute values. Even so, care must be used to reduce the noise picked up by the leads from the strain gauge to the amplifier that it drives, and to minimize a number of other error sources.

All these problems are solved by the optical power approach shown here, because the amplifier is located immediately adjacent to the strain gauge, and because the signal is then digitized before it is transmitted to the base station. In contrast with a conventional strain gauge application, no particular care need be taken with the installation of the remote amplifier, or the routing and shielding of the cable to the base station. It would therefore seem that there should be many applications for the optically powered strain gauge, even without the driver of the cost of high voltage insulation.

Some accelerometers are based on the principle of a strain gauge monitoring the deflection of a beam supporting a mass. Prompted by a suggestion from Ed Howells of Cooper Power Systems, we have done some preliminary work with accelerometers to study their application in monitoring the action of a circuit breaker.[2]

[2]In an ac power system, a circuit breaker is used to interrupt the current building up when a fault occurs. It is a matter of the physics involved that even though the contacts may move apart, an arc will be drawn until such time as there is a "natural" current zero. In a large circuit breaker, opening under fault conditions, the energy in the arc can be large. It is therefore usually a matter of policy to perform breaker maintenance following operation. However, many breaker failures follow maintenance.

In our laboratory setup, we used a small contactor and a 120-V circuit to simulate the circuit breaker, and slowed down its action. Since no effort was made to calibrate the accelerometer system, we have removed the labels from the vertical axes. The time axis is accurate.

In this test, the results clearly show that the contactor started to move near the peak of the current. The current was not interrupted at the first current zero, and an arc was drawn until the next current zero. The peak current is shown as reduced, presumably because of the volt-drop across the arc (which is a large fraction of 120 V).

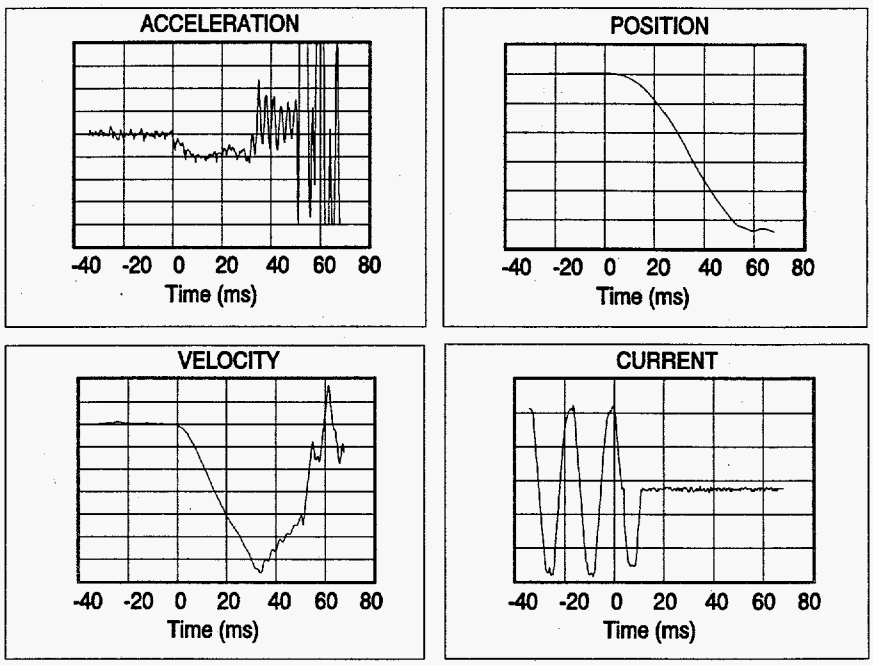


Figure 2-9. Results from a simulated breaker monitoring system. The acceleration data are integrated twice to show the position of the breaker when the current was interrupted

If the kind of information shown above could be stored by a local data acquisition system on the breaker, it could be downloaded later for analysis. It would provide the information needed to decide whether maintenance should be done.

However, it would be much simpler to obtain the necessary information by encoding the breaker position into its mechanism, and reading position directly. Readout could be done optically, even in the case of a live-tank breaker, at quite moderate cost. A data acquisition system monitoring the current and the breaker position could still make use of optical power technology.

TECHNOLOGY TRANSFER

In the last couple of years, these ideas were discussed with two manufacturers of electrical apparatus, with a view to technology transfer that would result in them being applied in the real world. First, we talked about a system to monitor the gas pressure in the interrupter chamber of an SF_6 breaker, part of a device called a *circuit switcher* by its manufacturer, S&C Electric, of Chicago, IL. The circuit switcher configuration consists of an isolator, or disconnect, and a

breaker in series, with only three insulators supporting them. Because the middle insulator, which contains the operating rod for both the disconnect and the breaker, is shared by the isolator and the breaker, the cost can be lower than the cost of a separate pair of switching components. However, the configuration does result in the use of a *live-tank* breaker, and monitoring the breaker requires moving information across the full line-to-ground voltage, without compromising the insulation system. The arrangement is shown in Figure 2-10.

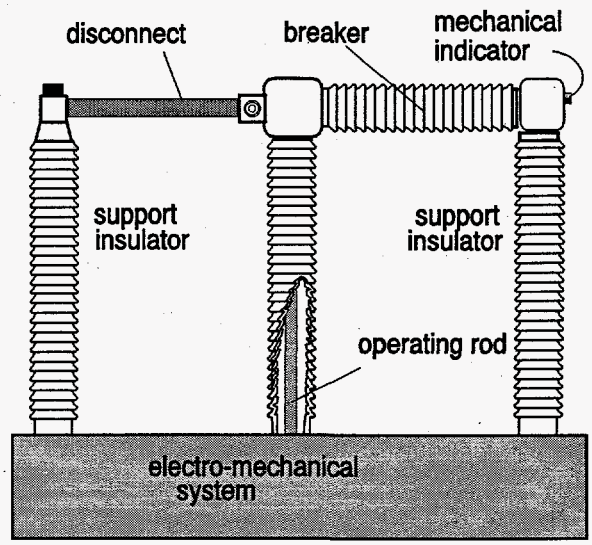


Figure 2-10. Circuit switcher, consisting of a disconnect (left) and a breaker (right) on only three insulators

While it is extremely unlikely, there is a possibility that the gas (SF_6) could leak from the breaker. This is a cause for concern both because of its impact on the operation of the breaker (the gas is used to suppress the internal arc when the breaker operates) and because of its impact on the environment (the EPA regards the gas as a threat to the ozone layer).

At the time that we started discussions with S&C, the circuit switcher was equipped with a mechanical indiation of low gas pressure, as shown in Figure 2-10. This was regarded as inadequate because the indication was not remotely telemetered, and system operators had no idea whether there was gas in the breaker.

In our proposal to S&C, we suggested that in addition to simply measuring pressure, a number of system improvements could be made. We had already demonstrated that an optically-powered system could readily measure a number of parameters, and we proposed to include a temperature measurement as well as a pressure measurement. There were three advantages.

First, by measuring the temperature at the point that the pressure was measured, we could correct for anticipated temperature-dependance of the pressure gauge. To meet the cost target for the S&C system, the pressure sensor would have to be inexpensive. Figure 2-11 shows the sort of interaction that might be expected from a low-cost sensor. This kind of nonlinearity could easily be corrected in software, either by a curve-fit of some simple algebraic function, or by a look-up table.

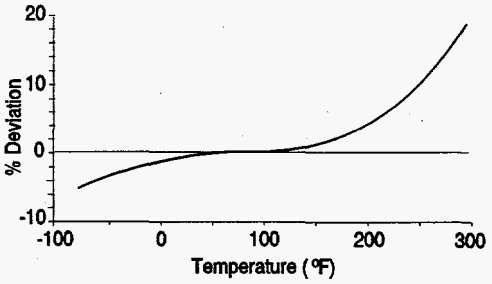


Figure 2-11. Possible temperature dependence for a low-cost sensor

Second, since there was a gas inside the brekaer, kowledge of the temperature as well as the pressure would allow a calculation of Boyle's law ($P \times V/T$ = const) so that decreases in the temperature that reulted in decreases in the pressure would not be mistaken for leaks.

Third, since the circuit switcher was an outdoor component, the temperature could become low enough that the gas began to condense inside the tank. Knowledge of the temperature would allow a comparison with the condensation temperature, so that this condition could be distinguished from a leak.

In the end, after several months of discussion, S&C decided to devlop a system that made the measurements we proposed, but communicated via a radio system. Power would come from a battery, rather than via an optical fiber.

Although the technology of low-power radio was also available to us, we had not included this in our proposal, as we thought the maintenance problem of battery monitoring and replacement made the optically powered system more attractive.

Another attempt at technology transfer was made. The potential sponsor was Lindsey Manufacturing, a company based in Azusa, near our lab in Pasadena, CA. For many years, Lindsey had been in the business of making low-energy measurement devices, aimed particularly at the low- and medium-voltage part of the power system.

For the JPL team, the appeal of the Lindsey approach was partly that it showed that the company was already convinced of the merit of low-energy sensing. The most common power system measurement is probably that of current, and this measurement is usually made with a transformer that is capable of delivering 5 A into its burden. The value of 5 A came about at the turn of the century, when a large current was needed to operate the relays. It had the further advantage that no fault was ever likely to induce a current of such magnitude into a protection circuit, so the chance of a false trip was small. However, as we approach the turn of the next century, the need for a 5-A signal to operate relays has all but disappeared (especially outside the US), and noise immunity is provided by other means (such as optics). Nonetheless, most CTs are capable of producing tens or hundreds of times the current actually used by modern relays. This is, most likely, a tribute to the industry's conservative nature. We had hopes that, since Lindsey was making low-energy sensors, we would encounter less conservatism.

Lindsey's transducers were typically housed inside line-support insulators, with low-energy signals being conducted by wire from the bottom. Electrically, the kind of system that was being made was similar to the basic arrangement of Figure 2-12, with a current transformer for the current and a resistive voltage divider for the potential.

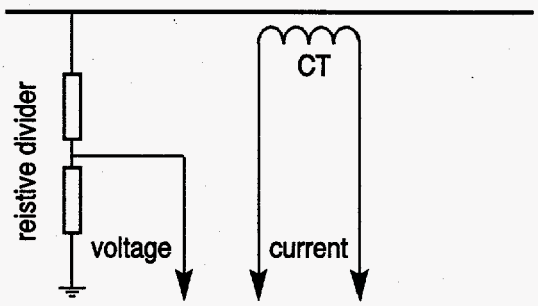


Figure 2-12. Low-energy current and voltage sensors for low voltage application

The proposal we made to Lindsey was essentially this. Turn the voltage divider upside down (so that the small-value resistor is closer to the line, and the large value one closer to the ground), and use a CT with no insulation. The money saved by having an uninsulated CT will *more than pay for* the additional hardware needed to encode the data and transmit it via fiber optic to ground potential. In other words, the product—based on our optically-powered electronics—could be sold competitively and profitably.

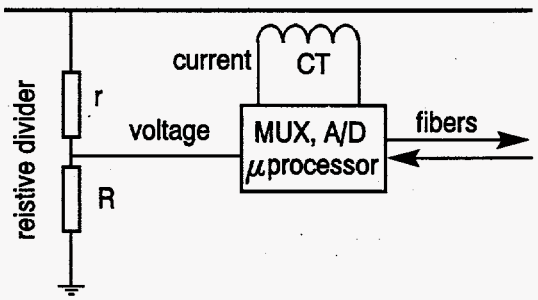


Figure 2-13. Optically powered version of low-energy current and voltage sensor

As further inducement to collaborate on a new product, we proposed a method whereby the temperature dependance of the components of the resistive dividers could be accommodated in software. In the existing scheme, the temperature dependenace of each individual resistor was measured by its manufacturer, and data were furnished to Lindsey describing this. Lindsey was then faced with sorting through batches of resistors looking for matches, so that the divider ratio would be less temperature sensitive, and the system would meet its specification. In our proposed improvement, the temperature dependence of each resistor would be plugged into the software, and the microprocessor would calculate the actual divide ration based on this information and an real-time measurement of the temperature. It was our opinion that the accuracy would be improved, and the manufacturing cost reduced.

We pointed out that with the coming dergulation of the power industry, there would be a need for measurements of this kind in many places where none had existed before: the new boundary between a distribution company and a trsnsmission company. (In the past, vertically integrated companies had owned both sides of this boundary, so no metering or billing was needed.) Therefore, the cost of developing a new product would soon be recovered by increased sales.

After months of discussion, during which we went so far as to lend Lindsey a prototype system, Lindsey decided that they liked the idea, but wanted to work with another supplier. As far as we know, all of our ideas are to be incorporated.

Perhaps there is a lesson here, though it may be a discouraging one. Both these attempts by the Communication and Control team to transfer the technology of optically-powered electronics failed because the potential sponsor went elsewhere. Our ideas were good, and indeed may have been included in the eventual implementations sought by S&C and by Lindsey. The basic idea of optically-powered electronics was not ours, was not new, and broad patent protection surely could not have been obtained. However, it is possible that narrower protection—while not, on its own, inerently valuable—might have allowed JPL to file a submarine patent. Submarine patents cover inventions that are unknown while the patent is pending. Expert patent submariners can keep a patent pending for years, or even decades, by filing changes from time to time. Eventually, it is hoped, some company will start manufacturing a product that infringes the patent.

One is forced to wonder whether JPL should start doing this, not so much to gain future royalties, but to change the outlook for a company that decides it can just take what it likes from government researchers.

CONCLUSIONS

We have shown how useful measurements can be made using optically powered electrical measuring systems. With careful design, the overall cost of the measurement need not be high. Even applications that are normally power hungry can be accommodated. The advantages of this approach are those common to optical measurements: immunity to noise, and reduction of common mode effects. In some applications, the higher cost of the measurement hardware may be offset by the reduced cost of implementing it.

Section 3: COMMUNICATIONS FOR THE DISTRIBUTION SYSTEM

3 COMMUNICATIONS for the DISTRIBUTION SYSTEM

3.1 INTRODUCTION

"Broadly speaking, an electric power system can be defined to include a generating, a transmission, and a distribution system. The distribution system, on a national average, is roughly equal in capital investment to the generation facilities. The sum of these two generally constitute over 80 per cent of the total system investment. Thus, it is readily seen that the distribution system rates high in economic importance, and represents an investment that makes careful engineering, planning, design, construction, and operation most worthwhile."

Thus begins the Westinghouse Electric Utility Engineering Reference Book on distribution systems, first published in 1959. The reality is, nevertheless, that the distribution system has not been the beneficiary of the careful engineering, planning or design that characterizes the other two parts of the utility, and its operation proceeds essentially unheeded and uncontrolled. The economics of the situation are that it has not been feasible to implement measurement or control in the distribution system to the degree possible in the generation or transmission system. The difficulty lies in the dispersed nature of the distribution system as compared with the generation or transmission system. There are hundreds of times more points to monitor, each handling perhaps only a few kilowatts.

It was a thesis of our work on communications for distribution automation, when we started in the 1980s, that there were still technical and economic impediments to automating the distribution system. We tried to overcome these impediments by developing using a communication system based on fiber optics, in an arrangement designed specifically for distribution automation.

This was done in the days when the utilities were protected monopolies, vertically integrated so that they included generation, transmission and distribution. Distribution was then the low-voltage dispersed end of a large power system. After the recent dregulation rule changes that have encouraged competition, some utilities are splitting horizontally, and concentrating on distribution, or generation, for example. Technically, the fiber-based communication system developed at JPL for the distribution system works whoever owns the network. In terms of its economic value, it may be that a robust distribution communication scheme will be more valuable to a company that specializes in distribution than to a vertically integrated monopoly. Time will tell.

This section of the Final Report begins by discussing the motivation for the work, and then describes the resulting communication system.

3.2 UTILITIES' INTERESTS

Both the utilities and the federal government were showing renewed interest in control of the distribution system when the JPL work started. Each had its own reasons. From the utilities' point of view, several factors stood out. First, as shown below, an increasing fraction of the total investment of the utility was in the least monitored and controlled part of its system. Planners were beginning to see better supervision of the distribution system as leading to better utilization of capital. Capital expenditures for system reinforcements could be deferred. Under some circumstances, the implementation of control in the distribution system could even result in the indefinite postponement of system reinforcement. Second, recent advances in communications and computing technology led system designers to the recognition that improvements were possible in their energy management systems. Better computers were available at lower costs than ever before. An improvement in the performance of an energy management system could translate into more efficient operation, and considerable economies. Third, distribution system control extensions are not usually large-scale efforts. The work can often be done piecemeal, and the capital required can be obtained in small increments. These various factors are still true, even for the deregulated industry entering the next century.

3.3 FEDERAL GOVERNMENT INTERESTS

From the federal government's point of view, there was concern about the long-term effects of the fall in oil prices in the wake of Saudi attempts to recapture their market share, and the collapse of the OPEC oil pricing structure in the mid-1980s. Statistics showed that the plunge in oil prices was leading to a degree of complacency about conservation on the part of the U.S. consumer. Oil consumption was increasing, and reliance on imports growing. Domestic output was lower at the end of the 1980s than its 1985 peak. Not the least of the problems is the effect of the imported oil on America's current account deficit.

The electric utilities, and industry in general, responding like other consumers to the low oil prices, increased their consumption of oil--by almost 18% in the year 1987 alone, according to an article in Science (Crawford, 1987). Apparently the situation had not improved in the decade since load management projects were examined under a U.S. Department of Energy program (Survey of, 1977). This report underscored the then government's concern over dependence on imported oil. Among the objectives of load management that were (and are) relevant to the national concerns listed above were shifting fuel dependency from limited to abundant energy sources and lowering reserve requirements. Both energy and capital could be (and can be) conserved.

3.4 MOTIVATION FOR RESEARCH

The conclusion that the United States will be increasingly dependent on oil (largely imported), natural gas and imported electricity seemed clear, and was disturbing to many. There was a possibility that our continued reliance on imported oil and electricity would lead to a larger current account deficit, and a weaker dollar, which would have to be propped up by higher interest rates. High interest rates hurt everyone, particularly U.S. industry.

These were issues that concern planners at the U.S. Department of Energy. Some of them were arguing for institutional changes[1], to restore to the utilities an environment in which economical base-load generation would again become possible. This would mean changes in rules at the state and federal level, including rate-making and construction financing. More efficient regulation, with fewer delays in construction, could save billions of dollars in carrying charges.

Meanwhile, it was recognized that the technical options had not been fully explored. It was hard to see how the efficiency of generation or transmission could be improved, but distribution systems had been the "poor relation" of the industry for so long that it seemed inconceivable that improvements could not be made there. Therefore the U.S. Department of Energy funded advances in the state of the art of distribution automation.

Other than for experimental or demonstration purposes, no utility has implemented a truly automated distribution system, yet. Nevertheless, some visionaries anticipate a highly automated distribution system by the end of the century (Caldwell *et al.*, 1982). Indeed, compared with other aspects of utility spending, distribution related investment seems set fair for the future. If this future is to include better operation and monitoring of equipment, it will require a better communication system than any now available in the distribution system.

The severe cost constraints that result from the large number of measurement and control points in the distribution system have meant that most of the distribution system is not monitored or controlled in a closed-loop manner. The addition of closed-loop control to an existing distribution system necessarily implies additional communications, either to the distribution substation or the

---

[1] "The utility industry has served the nation well for over a century. But the fact is that utilities are reluctant to undertake new construction to meet future needs. No new orders have been placed for nuclear plants since 1978, and there is little evidence of major base load construction involving other fuels. . . . If we are to avoid future capacity shortages, and to maintain reliable sources of supply, it is imperative that we look at ways of improving our current regulatory system and the incentives -- or disincentives -- we currently give to power producers." Martha O. Hesse, Chairman, Federal Energy Regulatory Commission, before the Electric Power '89 Conference, Washington, DC. Reported in DOE This Month, January 1989.

energy control center. The inadequacy of the available communications systems has been one of several impediments to the widespread use of distribution automation.

Another has surely been the fact that, for the most part, distribution automation has been a collection of functions, and not an integrated system. The list below makes clear why this could be. Distribution automation functions have become somewhat of a "wish list." Functions that are usually considered part of distribution automation may include:

management of customers' loads
monitoring of the performance of the power system itself
reading of customers' meters, perhaps even several times a day
detection of stolen energy
control of voltage in the power system
detection of outages
reconfiguration of the system following a fault
balancing of loads for optimal system operation
collecting load data for system planning

There are few engineers who would argue that distribution automation is a "system" and should be viewed as a complete entity in its own right. Most observers point out that several of the functions involved have no relation to one another, and accept as a working definition of distribution automation the implementation of two or more of the functions. The fact is, simply, that apart from a few trial installations, distribution automation does not exist, and so it may be defined very much according to one's own personal whim. If we accept that distribution automation is a system, then it is modular. The modules can be implemented independently and in stages.

Distribution automation has been extensively discussed in the literature, and the interested reader is able to find many IEEE Power Engineering Society (PES) papers and industry reports. Some of these are listed in a bibliography on distribution automation that covers the years 1969 through 1982, prepared by the Task Group on Distribution System Design (Buch et al., 1984), part of the Transmission and Distribution Committee's Distribution Subcommittee. The text of a tutorial course that has been offered at Power Engineering Society meetings can also be obtained from IEEE (IEEE 88EH0280-8-PWR).

3.5 COMMUNICATIONS FOR DISTRIBUTION AUTOMATION

There are a number of media that have been used for communication in the distribution system. None was ideally suited to all applications. Distribution line carrier (DLC) can be used to implement load management, or some such low-speed function. The data rate (typically a few bits

per second) is so low that messages must usually be broadcast, there being insufficient capacity to include addressing. Telephone can be used, but leased lines are point-to-point, and many lines would be required to reach all the places that distribution automation must access. The public telephone system can be used to reach consumers' premises or to communicate with distribution substations. In either application, the cost of service from the telephone company is likely to make the application difficult to justify economically. There are communications methods that use radio, but two-way communications requires two sets of equipment. In any case, there is a very limited amount of spectrum space available.

These technical and economic limitations, together with institutional issues such as ownership and the need for licenses, have been impediments to the implementation of distribution automation. Each time a utility wanted to add a distribution automation function, they had to consider a new communications system for it. They might, for example, have successfully automated the control and monitoring of a distribution feeder by means of a radio system, and then been required to add DLC to perform load management on the same feeder. Probably the radio system did not provide access to all the loads, or it may have been too expensive to modify the control software.

The result of this approach has been that electricity distribution communications has avoided a problem that faced most other users of communications systems. Most communications systems are able to support more than one user. In order to share the limited resource (the communications channel), the users have to abide by some agreed-to protocol. By and large, there is no need for such protocols in utility communications. Each user or application has its own dedicated communications system.

Ideally, the functions of distribution automation should each receive communication services as required. One might view the communication system as in Figure 3-1, as a cloud, with two-way traffic in and out of the cloud according to the needs of the function.

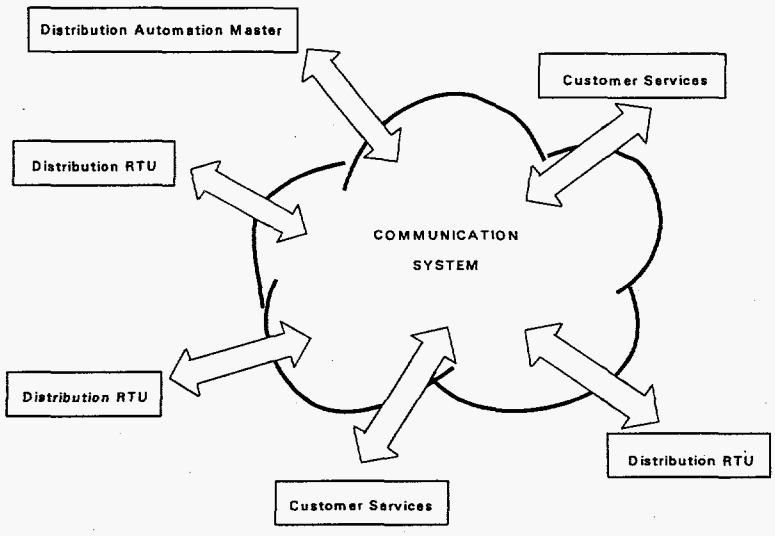


Figure 3-1. Distribution system communications viewed as an amorphous cloud

This is much more like the way computer communications networks have been obliged to address the problem. This view of things has the advantage that one is not obliged to worry over the details of communication. In the same way that the actual route taken by a telephone connection is of no concern to the caller, the details of communication should not concern the operator who needs to perform load management.

For distribution automation, communications may be needed anywhere in the power system. At a location where the power lines branch, there will likely be a disconnect or a fuse: these might require monitoring as part of a reconfiguration scheme. In a distribution line there might be a voltage control capacitor. While probably controlled open-loop by a time clock today, the system voltage would be better managed if some day the clock were replaced by a closed-loop control. All of these functions require an RTU to be located nearby. Figure 3-2 shows how the communication system is constrained to be topologically congruent with the power system.

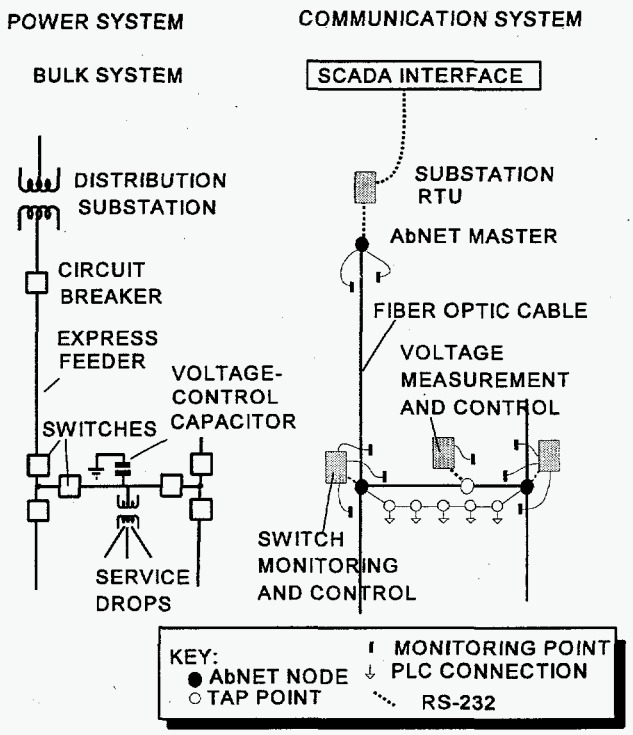


Figure 3-2. Comparison of distribution system and communication system topology

For control and monitoring of the power distribution network, a communication system that would meet the worst-case requirements is needed. This means that

- it can access as many locations as necessary to support monitoring or control functions,
- it can handle the highest data rate likely to be required by any foreseeable application,
- and it wil continue to operate even if part of the network were damaged.

Essentially, what is needed is a communication system that will be a transparent "phone company" for any and all automation functions. The topology of the communication network outside the substation (where control is assumed to originate), is congruent with the distribution system. This means that each time there is a lateral in the distribution system, for example, there must be a spur in the communication system. A master node would be located at each distribution substation. An example is shown in Figure 3-3.

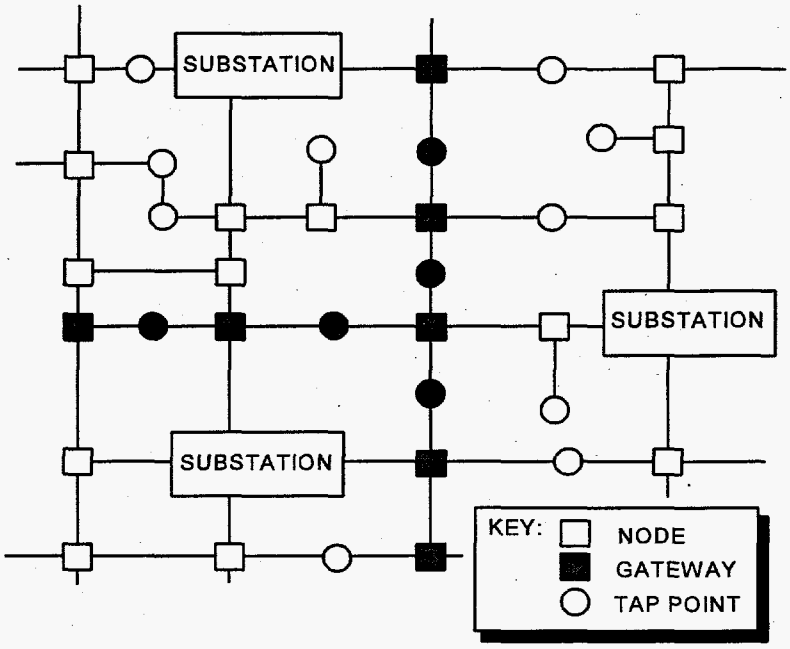


Figure 3-3. Typical AbNET system topology consists of multiple rings and loops

Normally, the distribution system is **operated** radially. There may be a handful of feeders from each substation, branching to serve the load. There is a limited set of loops that are normally open, but that can be reconfigured to provide an alternative way of bringing power to any given location. The fiber can, of course, cross a switch whether it is open or closed. As a result, the fiber optic communication system is arranged not as a conventional ring, star or bus system, but as a series of interconnected loops, with an occasional spur, as shown.

The power system requirements added to the uniqueness of the design, because the rest of the assumptions usually made in designing computer communication systems do not apply in the distribution system. There are two particular differences between LANs and distribution communications.

- Most local area networks (LANs) assume any user is likely to require communication with any other with equal probability, whereas the power system control application is hierarchical (master-slave), and peer-to-peer communication is rare.

- LAN topologies are usually fixed in advance, and are rather simple (loops or buses). In contrast, the configuration of a monitoring system for power distribution may have to change in the short term because of damage, and in the long term because the power system evolves.

Of course, the details of the communication protocols have to be solved, and the problem gets more complex as the number of users increases. Practically any physical medium can be used to allow just two computers to communicate. One may choose twisted pair, coaxial cable or fiber optics, depending on the speed requirements. The two users can agree on a data encoding method for the medium (baseband or carrier, signal levels etc.) and communications can begin. As soon as a third or fourth computer is added, additional questions have to be answered. If it is decided not to connect all machines to all other machines, and not to connect all machines to a common point, there are questions of routing and access to the common medium that cannot be left unanswered.

A series connection means that some messages must pass through an intermediate machine to reach their destination, but avoids the problem of contention for access to the common channel of a parallel connection. If a series connection is used, should the computers be arranged in a ring? Is a bus interconnection the right way to go? If the system uses a common communications medium, as in a bus, how is the access contention problem resolved? The addition of more users to the system brings more decisions that have to be made. A hierarchical structure, such as a tree network, might be useful. In a highly interconnected network, how are messages to be routed? As computer networks developed, the need for standards for interconnection became apparent. In this way, the location of other users and the manufacturer of their equipment could be made transparent to a user. The Open System Interconnection (OSI) model of the International Standards Organization (ISO) (Zimmermann, 1980) was a step towards standardization. This model split the problem of computer communications into seven logical and physical layers. While other arrangements are possible, and indeed were under study when the ISO model was released, the OSI model has become the accepted framework for discussing computer communications.

3.5.1 Open System Interconnection Model

Computer-to-computer communications in the ISO seven layer model is shown in Figure 3-2. It is important to realize that this OSI standard is just a model. Very few networks adhere strictly to this structure. Usually layers are missing because they are not needed in some particular application, and sometimes functions are implemented in layers other than the one to which this model would ascribe them.

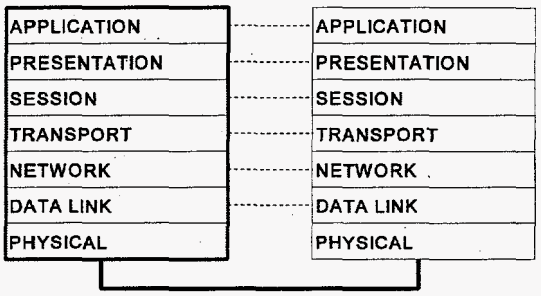


Figure 3-4. International Standards Organization model of open system interconnection

To make the model seem more familiar, we use the problem of getting two personal computers to talk to one another as an example. We begin this description with the top layer and proceed downward. Imagine that you have a personal computer with, say, a word-processing program. You want to send a file from your word processor to that of a colleague.

The **application layer** is the layer of software wanting to transfer data across the network. All other layers in the hierarchy exist solely to satisfy the needs of this layer. In the discussion example, this is the word processor program. In the case of a distribution automation system, the applications layer might be the SCADA program, the program that presents data to the operator, or the program that contains the control algorithms which are commonly part of SCADA operations.

The **presentation layer** provides services to the application layer to process the data in some way to make it more suitable for the layers below. This could mean, for example, translation or encryption of the data. The operating system of the personal computer is an example of a presentation layer. It furnishes data in a format that can be handled, for example by the word processing program or the screen driver, and at the same time can take a file of data furnished by a word processing program and present it for transmission to a remote computer to the layer underneath, the session layer. In the distribution automation example, this layer could contain user-callable library routines.

The **session layer** is the first of the layers in this hierarchy specifically concerned with communications to another computer. In essence this layer is responsible for coordinating interaction between the opposite end application processes. The layer has to be aware of both the application and the communications. Some of the functions of the typical communications program on a personal computer perform the functions required of the session layer. For example, a program sending blocks of data to make up a file might pause for acknowledgement at the end of each block. This could be a session layer function. In some applications, the session layer is a "virtual" layer. The decisions made here are the type of communication to be employed (e.g. full or half duplex) and how the failure of lower layers in the hierarchy is to be handled.

The session layer interfaces to the **transport layer** which is the highest of the seven layers responsible for the integrity of data. For example, this layer would be capable of performing error

checking, perhaps on a packet by packet basis. This function is performed in the personal computer world by a number of communication programs.

The job of the transport layer is to furnish error-free messages, in sequence, to the session layer. Whether this is a simple task or a complex one depends on the layers beneath. If the lower layers are retaining message sequence, and are performing error checking, the transport layer becomes very simple. On the other hand, if the lower layers can operate so as to get messages out of sequence, for example, the transport layer must correct the deficiency.

The transport layer connects to the **network layer**, which is responsible for furnishing data to the bare-bit manipulations of the data link layer below. In the PC world, the network layer is typically part of the telephone system, not the PC itself. The network layer is responsible for routing the information from computer to computer. Once you dial the phone number, you have no choice over the route that the phone company uses for your data. Consequently, it is their responsibility in this particular case to perform the functions of the network layer.

The **data link layer** is the highest level at which information as such is handled. The data link layer may perform error control on a word by word basis; for example, parity checking occurs in the data link layer. There may also be some means of flow control or hand-shaking to assure synchronization between devices capable of operating at different speeds. The telephone company for example, in their data link layer, typically will add bits to the 8-bit word coming from most PC modems, so that error checking (in addition to the parity bit that the user knows about) is usually performed.

Some users of the model have divided the data link layer into two sub-layers: **logical link control** and a **medium access control**. The logical link sub-layer is responsible for establishing, maintaining and terminating a logical connection between devices. The medium access control sub-layer ensures that only one device attempts to transmit at a time, performing the function of congestion control.

The **physical layer**, the lowest layer of this hierarchy, is the level at which the electrical signals are exchanged. A specification of the physical layer typically includes a description of electrical and mechanical quantities involved. For example, RS-232 describes the type of plug to be used as well as the voltage levels and the pin connections. In the world of modems, the physical layer includes a description of the procedures to establish or release connections between electrical circuits such as phone lines.

For most purposes in exchanging data from one computer to another, it is only the bottom three layers of this hierarchy that need to be standardized. However, because these layers do need to be standardized for effective communication between different machines, a good deal of work has gone on in this area. IEEE, for example, has developed a series of standards according to IEEE 802 Committee. These are usually known by the decimal organization of the subcommittees, for

example, IEEE 802.3 is CSMA/CD (Carrier Sense Multiple Access with Collision Detection) and IEEE 802.5 is Token Ring.

One may note also that the functional relationships between layers is clearly defined in the open system interconnection model, but is not so clearly defined in practice. The visibility of the boundaries between layers need only be clearly defined if the boundary corresponds to a product, in which case interfacing to that product (either physically or in terms of software) will be vastly simplified. Thus, the transport layer and the session layer functions are both performed adequately by PC communications programs; and the network layer, data link layer, and physical layer occur outside the boundaries of the typical modem.

3.5.2 Communications Issues for Distribution Automation

In arriving at the mesh configuration shown in Figure 3-3, we made the assumption that a single communications system would be used for all possible distribution automation applications. The system would therefore have to reach all parts of the distribution system, and be capable of handling all future control and monitoring data. The distribution system topology defined the communications network topology, and the data requirements led us to choose fiber optics as the communications medium. This, in turn, leads to the choice of a baseband system, probably using differential Manchester encoding. The need for multiple taps led to an arrangement of node-repeaters, rather than passive optical power splitters. While the system size (length of links, number of nodes) and topology are thus defined, there are some remaining system issues.

In a conventional communications system design, link data rate is determined by traffic requirements. For distribution automation, the data rate required is very low compared to the capability of the channel, and we can instead let cost considerations fix the data rate. While a determination must depend on the state of the technology, a bit rate of about 10 Mb/s seems to be achievable at minimum cost. Since the medium is fiber optics, the expected error rate can also be fixed by design. (Error rate is influenced by power margin and bit rate, but is unaffected by external factors.) As is common practice in fiber optics communications systems, a bit error rate (BER) of 10^{-9} is assumed.

This still leaves unanswered a number of questions of system operation. In terms of the seven-layer model, we have so far only defined the lowest. We have not addressed the question of how to avoid collisions, whether to use error checking to ensure reliable communications, how to route messages in the network, or whether voice and data can be transmitted over the same system. There are also questions of priority (should some users be assigned a higher priority than others?), and the general question of whether to adopt a distributed strategy for operation or a centralized one.

Generally speaking, these questions are addressed by considering the system application. A

telephone trunk application, for example, would indicate a design with high *efficiency*, in terms of the fraction of the time that the communications channel was carrying (revenue earning) data. Efficiency, in this sense, is affected by the channel bit rate (Mb/s), the packet length (assuming a packet-switching network) and the length of the trunk. A real-time application might be more concerned with *access delay*. Perhaps because it was not explicitly considered at the early stages of design, the access delay in some experimental distribution automation systems routinely amounts to tens of seconds, and is sometimes measured in minutes!

While transmission delay is likely to be important in many distribution automation functions, in our application the most important factor is likely to be *reliability* or *fault tolerance*. One of the goals of distribution automation is to improve the reliability of service by improving the performance of the distribution system. It must be presumed that a reliable communications system is a prerequisite. Ideally, the communications system should be failure-resistant, and preferably quite immune to single-contingency failures. The usual strategy for ensuring reliable communications at the lower levels of the hierarchy is error detection and request for retransmission. At the higher levels re-routing can be used. For the time being, assume that an effective error detection scheme is used to mitigate the rare occurrences of corrupted messages. What should happen if the network is changed by the operation of a bypass at a node, or by failure of a link? In other words, how should the network layer be used for maximum reliability?

If the substation unit maintains a map of the distribution system configuration, messages can be routed by a minimal number of intermediate nodes to any desired location. This would minimize transmission delay. Routing information for the response could even be included in the message. However, if the map is inaccurate (perhaps because of a line or node failure), this method breaks down, and some other strategy must be used. It may be possible for the central unit rapidly to update the map, but this approach seems cumbersome. We feel that any method that relies on a "retry" is inappropriate. Better to flood all paths in the communications network with information, than to use a more restrictive routing approach. It is possible that this approach may lead to lower efficiency, but it is our conviction that performance will still be within acceptable bounds for accomplishing distribution automation.

We may consider the communications network for distribution automation to consist of a number of interconnected rings in which information can be circulated. A conventional communications system with the stations connected in series would consist of only one ring. For this kind of network, communications protocols based on passing a token (which is a series of bits conferring the right to transmit) from station to station have been worked out. An example is ANSI/IEEE Standard 802.5-1985.

In this kind of system, a token is circulated in the ring, passing from unit to unit until one of them wishes to transmit information. When this occurs, that unit removes the token from the ring and replaces it with its message. At the end of this process, the token is placed back on the ring and circulated behind the message, so that other units can access the ring if they want to. In order to

prevent the endless circulation of the same information, the message is removed by an "active monitor" in the ring. This is a station that is specifically checking a particular bit of information (known as the monitor bit) in the system overhead data to solve this problem. The active monitor can remove a message from the ring, and thereby prevent its endless circulation.

3.5.3 Adapting a Token-Passing Ring Protocol: The Problems

The token passing strategy does not readily adapt to the multiply-interconnected rings or mesh configuration required for access to all portions of the communications network for power distribution. At least as of 1986, when our work on the problem began, there were no commercially available local networks based on the interconnected mesh (Hopper, Temple and Williamson, 1986, page 40). There are two problems:

First, the token can be duplicated in a system that includes branches. This can occur because, in order to maximize communications reliability, a node at an intersection of two rings is expected to insert any information it receives into both rings. Both rings might thus contain a copy of a valid token. Because of this, a message could be generated simultaneously by a node in each ring. At a point where two such rings come together, a collision (in the sense of simultaneous messages from two directions) would occur. Such a node must then include some means of buffering one message while another one is being processed. This is not a difficult problem to solve.

The second problem is that a single active monitor cannot prevent message circulation in a multiple ring system. Further, if the topology of the system is capable of changing during operation, it is impossible a priori to place one and only one active monitor in each ring. To solve this problem, each node must be able to terminate the process of repeating a message. A new, distributed protocol is required in order to solve this problem.

3.5.4 Distributed Network Control

Distributed strategies for solving the message-circulation problem have been discussed in the literature, and some have been implemented in practice. It seems, however, that no such strategy has been described for the multiple ring problem. In fact, there has been little work on multiple-ring networks, and what work has been done has assumed that the configuration is based on geographically small areas served by rings, interconnected and communicating only occasionally. Our system is rather different because its basic configuration is that of the power system. The multiple rings associated with one distribution substation are expected to communicate routinely to maximize communications reliability: occasional communication with an adjacent substation may also be necessary.

3.6 The Solution: AbNET

The problems of adapting a token ring system to mesh topology are mainly problems of the network layer of the hierarchy. Therefore, we describe the solution at this level first.

3.6.1 AbNET Network Layer

In order to effect a distributed strategy to solve the circulation problem, every message must contain a unique identifying number that can be stored in every node that it passes through. Since the operation of nodes in separate rings is assumed to be quite independent, the information required to identify a message uniquely could be based on the source address of the originating node. This address could be a "logical"address, or it could be simply the street address of the node. The identifier probably need only contain one or two additional bits, because it is unlikely that a greater number of separate messages could be simultaneously generated in any given ring. Alternatively, the unique identifier could be simply the source address and the time of origination of the message. The unique identifier numbers can readily be stored in a small "stack" in the node's memory. We will show later that the higher layers of the hierarchy provide us an even simpler way of identifying messages, but for now imagine that each message is somehow uniquely identified.

At all nodes, any message received is retransmitted on all outgoing lines, unless the message has been seen before, or is addressed to the node in question. By adopting such a strategy, a message inserted anywhere into the communications network shown in Figure 1-1 will be broadcast to all units in the network, and will not be repeated by a given node more than once. All links transmit the message exactly once. The method is therefore as economical in operation as a single ring with an active monitor. It is in some ways more efficient than a method which relies on some central unit working out a route for messages to go from node to node. The method, known as flooding, always chooses the shortest path between two nodes, because it chooses all paths. Transmission delay is therefore minimized.

Reliability is also maximized. The scheme will work even if a fiber is broken, and a node is therefore disconnected from one of its neighbors. (Such a failure can be detected by communications tests that could be ordered from time to time from the substation unit, but this is not an essential feature of normal operation.)

The communications protocol resembles the working of the body's immune system. The first time a T-cell in the immune system is exposed to an invading organism, it learns to recognize the organism as "non-self." On a second exposure, the T-cell will produce antibodies that kill the invader. In our communications network, the messages are the invading organism, and they wander throughout the network as far as available communications channels will allow. On first exposure to a message, our nodes store information that will allow them, on a second exposure,

to kill the message. Because of this similarity, we call the system "AbNET," after the microbiologists' abbreviation "Ab," for antibody.

3.6.2 AbNET Data Link Layer

Normally, the data link layer is responsible for delimiting data fields, acknowledgement of receipt of data, and some error control, such as parity check. In most communications systems, receipt of information that passes the error check is acknowledged to the sending station. The detection of an error will normally cause a request for retransmission. In addition to this kind of function, the data link layer may contain a flow control mechanism, to prevent problems when two devices of different speeds try to communicate. (One can hear this process when two modems exchange tones to establish a data rate.)

Some simplifications are possible in AbNET. First, the bit rate of all nodes can be fixed in advance, so flow control should not be needed. A data encoding scheme such as differential Manchester can be used, so that receiving devices can synchronize themselves to the incoming data, and so that there is no net dc level in the signal (an important consideration with optical communications).

There seems to be little to be gained by retaining the conventional technique of message receipt acknowledgment. With a BER of 10^{-9}, the probability of sending a message M bits long without error is $P = (1-10^{-9})^M$. If the message is 1000 bits long, the success probability is 99.99990%. In other words, only one message in a million would require retransmission or correction. It seems inherently wasteful to devote any amount of effort to acknowledging successful reception of data. We use instead an error detecting code that reduces the occurrence of uncorrected errors to insignificance.

For now, note that the absence of a retransmission-request procedure means that all traffic in the network is associated with the first-time sending of messages. No link will be made busy by nodes trying to overcome neighbor-to-neighbor communications problems, so it can be guaranteed that there will be no network congestion.

3.6.3 AbNET Physical Layer: A Hybrid System

In the discussion above, it was assumed that the communications system was exclusively fiber optics. While this is technically feasible, there are considerations that make other communications media worth considering.

First, in a distribution system that is widespread, such as a rural network, the economics do not favor the use of a fiber optics medium. As we showed earlier (Kirkham et al., 1989) UHF radio

can be cost competitive at node spacings above about 500 m. Radio transceivers operating at about 950 MHz are costly items, but if their cost is less than the cost of the fiber cable required to cover the distance, their use may be justified. The greatly decreased speed (bandwidth) of the radio system may be a problem, but if the application can be adequately supported even with the reduced performance, radio may be the appropriate choice.

There is nothing in the protocols that would make AbNET unsuitable for a radio-based system. In essence, the fiber optics network considered earlier is a broadcast system, with all nodes receiving all information. A radio system is different in that a receiver could conceivably receive a signal both direct from the base station and repeated by another transceiver. As far as the AbNET protocols are concerned, this would be no different than a node receiving a message over two different fibers. The second message would be ignored.

The second aspect of the physical layer that makes non-fiber systems worth consideration is access to customers' premises. The use of a fiber optics connection for load management, for example, would mean access to the home, as well as modification of existing designs of load management equipment. A hybrid system, (and by this we mean a true hybrid system, that requires no systems engineering to be done by the utility--a marked contrast with today's combined systems) could overcome this disadvantage.

A hybrid that used a high speed network on the distribution feeders (required by feeder automation functions) and low-speed communications on the secondary is attractive. An example of this kind of system (actually UHF radio and DLC) is used in the NetComm project in California (Holte, 1989). We propose to specify AbNET as a hybrid system consisting of fiber optics (or possibly UHF radio) on the distribution feeder and DLC on the secondary.

The use of DLC on the secondary has several advantages. This approach enables the continued use of existing load-control and meter-reading hardware, which in turn implies minimum need to enter the customer's premises for installation. Further, communication into the home can be compatible with other communication applications of the home wiring, for example the PLBus of the Electronic Industries Association (EIA) Home Automation Standard. An EIA PLBus interface at each distribution transformer would allow the utilities to accomplish far more control than simple management of one load. Not only could there be the capability of emergency load-shedding, but a solution to cold load pickup problems would be available.

In cost terms, a hybrid of this kind is certain to be competitive if more than one function is to be performed by the distribution automation system. Feeder automation functions can be implemented competitively (see above) and demand side management can be added at extremely low cost. The small size of the network to be covered by the DLC or home wiring system (typically 3 or 4 houses) greatly eases its communications requirements. Very low power, inexpensive modules operating at about 120 kHz are widely available, and signalling speeds up to 1200 baud are possible. The incremental cost of adding this capability and other functions at the customer level

would be minimized.

Complete physical layer specifications will be addressed in our future work.

3.6.4 AbNET Communications Management: The Higher Layers

In computer communications, it is generally assumed that each location is as likely as any other to originate a message. This assumption leads to the need to solve the contention problem for access to the medium. A similar assumption has generally been made the past in designing communications systems for distribution automation. Because of the generally low data rate of available communications channels, it has been necessary to limit communications traffic by having the RTUs make some intelligent decisions on their own. RTUs are usually designed to operate with software that permits them to originate a transmission only if they detect some drastic change in the data they monitor.

Since the cause of such a change cannot be determined in advance, and might affect several RTUs simultaneously, some kind of medium access protocol must still be used. Collision detection has been used in some demonstrations, and is an automatic feature of the busy signal of telephone-based communications. The problem with this solution is that the response to a collision is always a delay. Whenever a large or widespread change occurs in the distribution system, the communications system becomes overloaded, and information transfer is subject to unpredictable delays.

This is not the case with AbNET. A centralized polling strategy will be used, so that a node (RTU) can transmit only if so directed by the substation unit. In essence this is a token-based system, but only the substation unit can originate a token, and that is addressed to a particular RTU. Before the reader objects that this is slow and inefficient, since it requires that all nodes be scanned before a problem can be discovered, let us point out that we expect a scan of all nodes to occupy less than 20 ms. If the message from one RTU is 1000 bits long, and there are 100 RTUs, a complete scan comprises 10^5 bits. At a data rate of 10^7 b/s, the minimum scan time is 10^{-2} s, or 10 ms. We have doubled this estimate to allow for error correction code overhead, and software execution time.

An advantage of the polling approach is that it allows the substation unit to allocate communications time dynamically. If a problem is suspected in some part of the system, the central unit can concentrate on that area, and relegate the remainder to a lower priority. Whether this would ever be necessary will presumably depend on the application.

It is also apparent that in a polling system, the central unit can simply assign message numbers to its transactions with the RTUs. The first byte of the data could be used as the message identifier, for example. A responding RTU could simply increment the message number. There would then be no need for the RTUs to keep track of the time, or the originator of a message they were handling.

The widespread use of optical bypasses in the network means that the substation unit (the central unit for any given feeder) must be capable of updating its view of the system from time to time. Because of the interconnected nature of the communications network, failure of a single node is unlikely ever to prevent communications to any other location. Such a failure may even be relatively unimportant to the control or monitoring functions for a long time. Nevertheless, the operation of its bypass should not go unnoticed or unrecorded. Because the communications system management software in the central unit can address all nodes, it can be made capable of developing its own system map when it is first turned on.

Suppose each remote unit is equipped with a ROM describing its location, and a few pertinent details of its function. At switch-on, and periodically thereafter, the central unit can poll all the remote units to determine their status and their connectedness. This might be done as frequently as every second. Any changes (in the communications system, not the power system) could be logged automatically. Software could be written so that the operator display could be automatically updated to reflect problems such as the inability to communicate with a particular location.

Even failure of a substation unit can be handled without overloading the communications channel. In the system diagram Figure 3-3) some nodes were designated "gateways." A gateway is a node that can pass signals from the area of one substation to another if required, so that communications can keep pace with the changing power system configuration that might follow loss of power into a distribution substation.

3.7 Modifications to Original AbNET Design

AbNET is a dual-hybrid network, comprised of a fiber-optics-and-power-line-carrier digital data system in which every node (or RTU) is also a repeater. A passive fiber optics PCM (pulse code modulated) voice channel shares the fiber cable. The nodes are interconnected by fiber optic links routed along the distribution system. The fixed installation of the voice channel is intended to be passive, the optoelectronics and the electronics being carried by the user and inserted into the system as needed. Only the digital data system can be put in the framework of the ISO OSI model. This is shown in Figure 2-2.

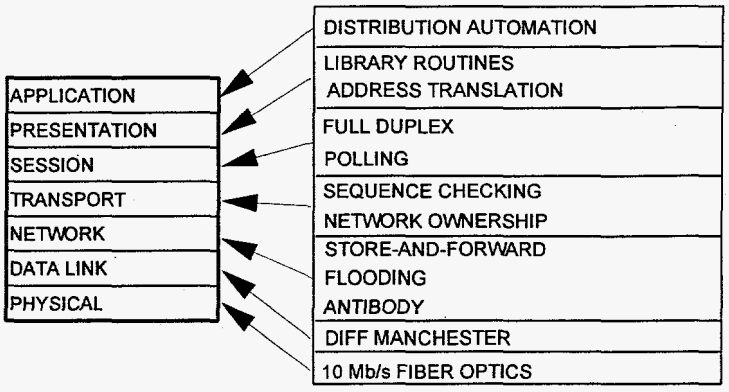


Figure 3-5
Main features of AbNET
in OSI framework

- Distribution automation is shown in Figure 3-5 as the applications layer. The bulk of the routine traffic will probably be associated with data acquisition. Descriptions of the various other functions that comprise distribution automation need not be repeated here. Suffice it to say that AbNET was developed explicitly to satisfy the communications requirements of data acquisition and control for distribution automation. The brief descriptions that follow show how AbNET meets these requirements, and supports distribution automation.
- The principal function of the presentation layer is to perform the conversion of data descriptions (in application programs) into addresses that the lower layers can recognize. •

The session layer establishes transport connections between the central unit and the remote nodes on a polling basis. Centralized information about network performance can be made available to an applications program only through the operation of this layer. The session layer therefore retains the ability to establish (and test) connections between adjacent nodes. Since message delivery can occur in the AbNET system even if a number of connections are broken, network integrity is checked by software, operating transparently to the user, at this level.

- The transport layer is centralized at the distribution substation. It normally operates through the local network in support of the data acquisition and control functions. Under unusual circumstances, it can choose to use the local network for access to the area served by another substation.

Normally, part of the packet header is the node address, a 2-byte logical designation. One of the two bytes is used in AbNET to indicate "ownership." A substation owns all the RTUs in its

service territory, and none belonging to the neighboring territory. Ordinarily, network layer operation screens the higher layers from messages intended for RTUs in other areas. Along the border of two service areas, however, the RTUs (called gateways) have to respond to two (or more) master stations at the substations. This means that the rejection of "foreign" messages occurs not at the boundary itself, but one layer of RTUs deep into each service area.

This feature provides a mechanism for one master station to take over the territory of another if it is ever needed. A special kind of message can be sent from one master to the RTUs of another, re-designating ownership, and redefining the boundary.

- The network layer is a decentralized message-based, store-and-forward system. Congestion in the network is controlled by recognition of ownership, and by an antibodylike algorithm, that allows a node to repeat a message once and only once. Routing is accomplished by network flooding.

- The data link layer incorporates the field descriptions and framing definitions of IEEE 802[2] as far as possible. Because the bit error rate is expected to be extremely small, there is no handshaking for flow control, and no request for retransmission in the event of an error. This avoids system problems in the case of hardware failures.

- The physical layer is a fiber optics based hybrid. Multimode fiber can be used, transmitting optics will be based on LEDs operating in the near infra-red, the bit rate will be 10 MHz, and differential Manchester coding will be used. (In many respects this is similar to Ethernet). Access to locations on the power line secondary will be by means of low-power high-frequency power line carrier communications.

6.0.1 Additions to AbNET Design

In operation, the application program performs its function by passing instructions through the presentation layer to the session layer, for coordination with the communication system. The session layer is responsible for polling the RTUs, and gathering data from them. Although it was never stated explicitly, it has been generally assumed that the polling would be sequential, with

---

[2]For most purposes, and in the case of most standards, it is only the bottom three layers of this hierarchy that are standardized. However, because these layers do need to be standardized for effective communication between different machines, a good deal of work has gone on in this area. IEEE in particular has developed a series of standards according to IEEE Computer Society Project 802. These are usually known by the decimal organization of the subcommittees. For example, Ethernet has been adopted as IEEE 802.3 or CSMA/CD (Carrier Sense Multiple Access with Collision Detection) and IEEE 802.5 is an embodiment of IBM's Token Ring.

each RTU being addressed in turn. This would follow naturally from a program that expected data from each RTU in turn. While this will work, and with the AbNET physical layer it is fast enough to meet the requirements of SCADA, a speed improvement is possible at the cost of a slight decrease in reliability. This modification permits the application layer to issue an instruction equivalent to "Poll all nodes" to the lower layers.

Consider the section of a network shown in Figure 3-6. A master station has control of some RTUs that happen to be connected in a loop. If each is addressed in turn, the polling sequence would be: Poll 1 (wait for response); poll 2 (wait for response); poll 3 . . . and so on.

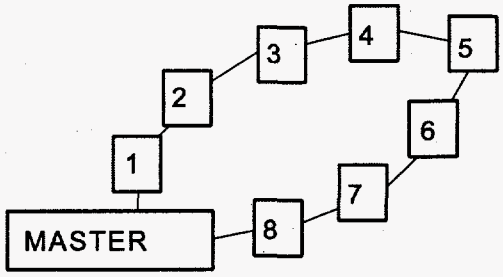


Figure 3-6
Example of simple network

The message handling sequence is more complex, however. Thus, the sequence begins: Master to 1; 1 to Master; Master to 1, 1 to 2; 2 to 1, 1 to Master; just to poll the first two RTUs. As the length of a line (in terms of the number of RTUs on it) increases, the number of "hops" required to poll it increases faster. Considering a linear network, the number of hops required to poll the RTUs is as shown in Table 3-1.

Table 3-1
Number of hops required to poll a linear network

nodes	hops
1	2
2	6
3	12
4	20
5	30
6	42

In a linear network, such as the one shown in Figure 2-3 if the connection from the master to node 8 is broken, the distance from the master to node j is j lines. To poll this node, and to obtain the response, therefore takes $2j$ hops. The number of hops H required to poll all the nodes up to node n is therefore simply

$$H = 2\sum_{j=1}^{n} j$$

It can be shown[3] that this expression reduces to $n(n + 1)$, which greatly simplifies evaluation. In the network of Figure 2-3 the most distant node is only 4 lines from the master. The network will therefore be polled in 20 hops. In a system that has 10 lines between the master station and the most distant node, it will take 110 hops for the last response to be received.[4]

The total number of hops for a data acquisition scan is reduced if the flooding algorithm of the network layer is modified. In the general flooding network, a node acting as a repeater retransmits an incoming message on all outgoing lines. Similarly, it transmits its response to a message addressed to it on all outbound lines. In the network of Figure 2-3, for example, node 2 responds to the master on both the line to node 1 and the line to node 3. Unless there is a network failure, only the response to node 1 is important.

Taking advantage of this, suppose the action at node 2 consists of a response to the master, on the line to node 1, and a simultaneous repeat of the poll on the line to node 3. This change has the effect of reducing the number of hops required to poll the network to the number required to poll the most distant node. The system of Figure 2-3 can be polled in 8 hops; a 10-line system can be polled in 20 hops. The improvement due to the modified flooding algorithm for these two cases is thus between a factor of 2 and a factor of 5.

There is a drawback, but it is relatively minor. If the network configuration changes between the reception of the poll and the end of the response, messages could be lost. Suppose the line between the master and node 1 fails immediately after node 1 receives the request for data. In this

---

[3]By definition

$$\sum_{j=1}^{n} j = 1 + 2 + 3 + \cdots + (n-2) + (n-1) + n$$

reversing the order:

$$\sum_{j=1}^{n} j = n + (n-1) + (n-2) + \cdots + 3 + 2 + 1$$

adding term by term:

$$2\sum_{j=1}^{n} j = (n+1) + (n+1) + (n+1) + \cdots + (n+1) + (n+1) + (n+1)$$

There are n terms in the series, so that:

$$2\sum_{j=1}^{n} j = n(n+1)$$

[4]Even this large number does not result in unacceptable delays. If each individual message is 1000 bits long, it takes 100 μs to insert the message into the network. With no allowance for processing, 110 hops thus takes 11 ms.

case, the responses from nodes 1, 2, 3 etc never arrive at the master. Redundant copies are not circulated in the other direction. While a copy of the poll might arrive at the low-numbered nodes from the other direction, these nodes will judge that they have already responded. However, within a few ms the master can detect that node 1 has "timed-out," and that several of the nodes have not responded. If a new poll is generated (and if the network does not change again), the missing data can be obtained, the poll being routed automatically by the flooding algorithm.

The loss in reliability is relatively minor because data are lost only when the network configuration changes. Once the configuration has stabilized, the two versions flooding behave identically. Field testing of the two versions may indicate the relative value of the performance improvement and the reliability loss.

Another, related, concept worth considering in the context of squeezing the maximum performance out of the AbNET protocols is "rapid-fire" polling. In a full-duplex system, there is no need for the master station to wait for a response before transmitting a new poll. One can visualize a network carrying an endless train of polling messages away from the master, and a more or less continuous stream of data returning in reply. In the original polling scheme, the master station waited for each node to respond before the next was polled. In the revised polling scheme, the master station waited for a response from the set of all nodes before starting a new scan. It is difficult to estimate the speed improvement that would result from waiting only for the output channel to become clear before sending out a new poll. An assessment may be made in the near future using the software simulator (described below).

One further refinement is worth mentioning. In a practical distribution automation scheme, there are likely to be many RTUs dedicated to customer interaction, such as meter reading or load management. These RTUs are nodes that can be expected to be connected in a linear fashion along the feeders, contributing significantly to the number of nodes to be polled in a single scan. In terms of system operation, such RTUs will likely have little data of value. The AbNET protocols therefore recognize different classes or priorities of RTU. During normal SCADA operation, the low-priority nodes will act as repeaters, adding slightly to the message transit time. They will not furnish data unless they are specifically addressed.

In the original flooding scheme, not polling low-priority nodes could lead to significant reductions in the scan time of the system. In the modified protocol, there would be no effect other than a reduction in the amount of data delivered to the master.

6.1 Software Simulator

In order to make a convincing demonstration of the operation of the network protocols, a representative topology was planned. The network should have a large number of intersecting

loops or meshes, and a small number of nodes that were served by spurs. Since the network was supposed to represent part of the distribution network served by a substation, it should have more than one line from the master station. While the underlying power system could be assumed to be operated radially, the communication system could bridge open disconnects. In collaboration with George Allen of American Electric Power Service Corporation,[5] the configuration shown in Figure 3-7 was decided upon.

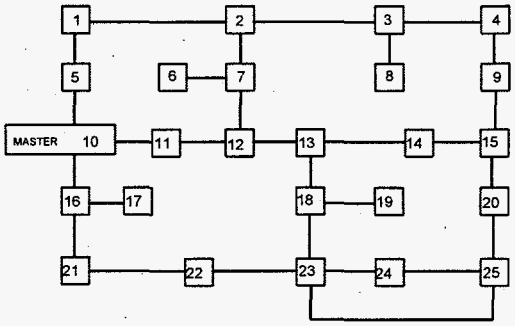


Figure 3-7
Representative portion of communication network

In order to show the flexibility of the network, a variation with two master stations was also designed. This is shown in Figure 3-8.

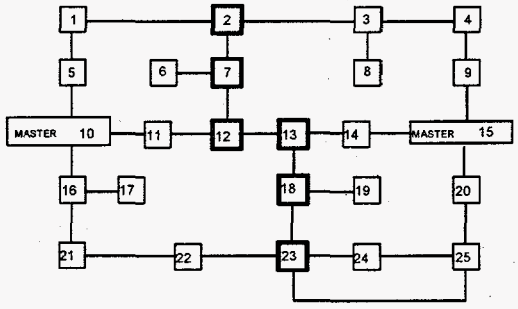


Figure 3-8
2-Master version of communication network

In this version, nodes numbered 2, 7, 12, 13, 18 and 23 are of the type we have designated "gateways." In the AbNET system, a gateway differs from an ordinary node in that it responds to two (or more) masters. In the Ethernet terminology, it is "promiscuous." Gateway nodes are located along the border between two service territories. Rather than put two RTUs on a pole at a border, each owned by one particular substation, the gateway RTU is owned by both substations. Not only does this make economic and aesthetic sense, it adds the flexibility to move the boundary by means of commands sent over the network itself. There are a number of quite realistic scenarios in which such a feature could be advantageous in distribution automation.

---

[5]At the time, AEP was considering a "real-world" demonstration of the AbNET network, to be held at their John Dolan Laboratory near Columbus, Ohio.

6.1.1 Overview

The original purpose of the software simulation (as opposed to the demonstration in Ohio) of this network was to investigate the impact of operating a half-duplex protocol. While it had been intended to operate full-duplex from the very beginning, the laboratory scale demonstration, using modified Ethernet hardware and ordinary PCs, could be done more easily and at lower cost if the protocol were half duplex. The question to be answered was How much impact would half-duplex operation have on system performance. Accordingly, a simulation was written in which the network of Figure 2-4 and the networks of Figure 2-5 were modeled in considerable detail.

The simulation was written in QuickBASIC, for convenience and portability.[6] The three major parts of the program will be described.

- First, the initialization section, executing once, is used to determine which network to simulate. There are a number of other user options. Since the simulation uses only one processor to emulate the behavior of the 25 processors in each network, it was necessary to examine each in turn, to check for messages, generate responses and so on. Because this meant a deterministic approach to the performance of the system, the option was included to have the master stations generate destination addresses either in sequence, or at random. The option is also offered to take lines out of service, or restore lines to service that have been removed. After initialization, a message is created and stored at the master (or masters), and a flag is set to indicate to the next part of the program that a transmission is ready.

- The remainder of the code operates iteratively. The simulation portion begins by examining each node to see if it has a message to send. Flags have been set (either by the initialization section, or by the previous iteration) for this purpose. If a transmit flag is set for a node, the message it has to send, stored in a transmit buffer, is copied to the connected nodes, unless a flag derived from considerations of full or half duplex blocks the copying. This is done for all nodes.

Next, the program scans each node again to determine what action it would take operating under the control of the AbNET protocols. For example, if a node has received a message, it checks to see whether it should respond, repeat the message, or ignore it. At this point, a new set of transmit flags is set, and the iteration repeats.

- As the simulation executes, a display portion is called to show the operation of the protocols. A map of the network is shown on the screen, with the nodes numbered, and with

---

[6] The simulation consists of about 4000 lines of code. The complete source code will not be published in this report. Readers who are interested in experimenting with the simulation, may obtain a copy by sending a DOS-compatible disk to the author, within a year of the publication date on the cover.

downed lines indicated. Each message is associated with a color, and the color corresponding to the last message handled at any node is used to fill the box representing the node. In addition, a vertical bar showing the source and destination of the last eight messages, along with their color, is displayed for each master station. A tally is kept of the number of nodes that have responded to the poll.

6.1.2 Results

The original purpose of the simulation was to see whether the performance of the system would be adversely affected by half-duplex operation. At the time of writing, only the original AbNET protocols (the ones designed prior to the modifications described in Section 2.1.3) have been simulated.

One of the first results of the simulation was that, in the single-master network, exactly the same number of nodes could be scanned in a given time in the half-duplex version as the full-duplex version. In other words, there were **never** any delays in transmitting messages because of having to wait for traffic in the other direction.

This may at first seem surprising. The result arises from the combination of polling and flooding. This combination means that no RTU can spontaneously originate a message. Messages flood out from the master, and when a copy reaches its destination, the response floods back. There is a possibility of two messages arriving at an RTU at the same time, but they will arrive over different fibers. This is permitted, and does not result in a loss of data.

There is a small defect in the half-duplex version of the model that should be mentioned. Because each node is examined in turn, low-numbered nodes are allowed to send, and the higher-numbered nodes they send to are made to wait. This is somewhat unrealistic. In a real network, there would arise some occasions when the two nodes attempted to send to one another at the same time. In the absence of handshaking, neither message would arrive.

In the example of Figure 2-3, nodes 4 and 5 would attempt to send a copy of the polling message to each other at the same time. While this would mean that neither message arrived, it is also clear that there is no loss of information, since both nodes have already handled the poll in any case. It is generally true that a collision of this sort, concerning identical data, will not result in the loss of information. It can only occur over links that happen to connect two parallel paths.

On the other hand, an overlap on a link could happen in the case of multiple masters, where the messages in the area of the gateways are different, but are ready to be sent at the same time. Consequently, it can be argued that the results of the model are optimistic in the case of two masters, since the code allows messages to get through that should not. In defense of the simulation, it can be said that it contains a compensating bias towards pessimistic results in the

two-master case. In the model, an unrealistically large number of the RTUs are gateways. The model gateways are therefore obliged to handle a larger fraction of the network's total traffic than is realistic. Because the program examines low-numbered nodes first, high-numbered nodes can be made to wait. The effect of message overlaps is therefore exaggerated.

The model shows that half-duplex operation is not more than 10% slower than full-duplex, for the 2-master network of Figure 2-5. Since this network has the high proportion of 6 gateways for 17 ordinary nodes, it is reasonable to expect that a realistic network would perform rather better.

Of greater interest to the operator is the demonstration by the simulation of the automatic routing feature of the protocols. Arbitrarily chosen lines can be taken out of service, and new routes to the destination are found. With colors representing messages, this becomes a fascinating exercise.

6.2 ZONES

At the border between two service areas are gateway nodes. In most communication systems, a gateway is a node that can pass signals from the area of one communication system to another if required. Often, such a gateway will examine the destination address of a message, to see where to send it. In the AbNET system, such selectivity is reserved for the ordinary nodes. These nodes will only handle traffic associated with their master. The gateway nodes operate in what ethernet calls the "promiscuous" mode. They will handle any message presented to them. As a result of this approach, the nodes that are immediately adjacent to the border see additional traffic generated on the other side.

This reversal of normal roles has important advantages in the power system. First, at the border between two zones, or even two companies, only one communication node and one RTU need be installed, instead of two. Second, communications can keep pace with the changing power system configuration. Special messages can be sent from one zone into another to redefine ownership. Should power be lost to a substation unit, control of its service area can be recovered, and assigned to neighboring substations.

6.3 SERVICES

A feature of a highly fault-tolerant network is that failure of any particular connection may not be detected, and several failures may have occurred before there is a communication failure. There are a number of ways this drawback can be overcome; one of the more obvious ones is to have intermediate nodes update a routing table in the message header. Changes in the route taken by messages can then be logged. This approach increases packet size, and may not detect changes in parts of the network that are redundant, because duplicate copies of packets are dropped.

An alternative is to temporarily suspend the distributed flooding algorithm, and instead use a part of the network layer that is designed specifically to have each node identify its nearest neighbor.

The remote nodes build their own table of neighbors, which they send to the master periodically. Because the message flow required for this task resembles the operation of Sonar, we have designated this a Ping.

This approach takes time out of normal network operation. In the prototype network, a ping operation takes place about every 10 seconds, and a map at the master station is updated at this frequency.

6.4 SYSTEM SPEED

The requirements of the power system application dictated that the communication system be able to support data acquisition from a typical network at a rate of one complete scan every three seconds. We have estimated that this amounted to a data rate of about 3 kb/s on a feeder (Kirkham, Johnston and Allen, 1994), which on its face presents no challenge to a fiber based network.

However, while the fiber network may be underutilized at such a low data rate, the overall system may be challenged to meet the requirement. Suppose that a request for data is sent from the master station to the first node, and that the master station waits for a response before continuing. In the power system application, the RTU will be connected to the communications node via an RS-232 connection. If this is operated at 9600 b/s, and if the return data occupy 100 bytes, there will be a lag of about 100 ms before the polling sequence continues. This would limit the network size to approximately 30 RTUs. While this may be acceptable in many applications, it was thought that it should be possible to support a larger number. The polling approach was therefore modified. At present, four varieties of polling are available, as shown in Table 3-2.

AbNET 1 is a polling approach in which the master waits for a response before moving on. In AbNET 2 and 3, the polling message is broadcast, and the slower RS-232 communications between the communication nodes and the RTUs take place at the same time throughout the network. (The difference between variants 2 and 3 is a small change in the sequence of events at the node after a poll command is received. AbNET 2 would immediately begin to pass the poll to the RTU, and repeat the poll message onward, whereas AbNET 3 would immediately flood the poll in all directions. In practice, the distinction may never be noticed.)

Table 3-2. Comparison of protocol variants

Protocol type	Polling message		Response Message	
	Addressing mode	First hop relay method	First hop direction	Timing
AbNET 1	Individual	Flood	Flood	
AbNET 2	Broadcast	Onward only	Return path only	Simultaneous
AbNET 3	Broadcast	Flood	Flood	Repeat poll first
AbNET 4	Groups	Flood	Flood	Bursty

The AbNET 4 variant allows the master station to serve several groups of RTUs, and to distinguish between them. In this, it is the one variant that does not maintain complete transparency of use. It has the advantage that it can allow several "virtual networks" to co-exist on the same physical network, without interacting. This means that until there is complete interoperability of RTU protocols, RTUs with different protocols can safely be operated together. This variant was designed by Licom, the licensee of the AbNET system for power system applications.

With any of the variants of the original protocol, the time taken to poll the network for data is approximately the same as the time taken to poll a single RTU. While the amount of data in the system at any time is greater than in AbNET 1, it is still small by fiber standards. The maximum utilization factor is roughly the ratio of the RS-232 speed to the fiber speed, multiplied by the number of nodes. With 100 nodes running continuously, the fiber system would be loaded at about 50% capacity. With the normal 3-second scan, and with the anticipated amount of data, the fiber system would be loaded to less than 1% capacity. The spare capacity may find future use for other functions.

6.5 RELIABILITY

The physical layer is assumed to have a bit error rate of less than 10^{-9} for each link. With an average packet size of 100 bytes (approximately 1000 bits), about one packet in a million is corrupted and without error correction will be dropped by the receiving node. In a system of 100 nodes, scanned every 3 s, there are about 10^9 packets per year. Since the typical packet travels more than 1 hop, without redundancy more than 1000 packets per year will be dropped. (To avoid data loss, the SCADA system would re-poll if slaves fail to respond within some timeout period, essentially performing a transport layer function.)

For a packet transmission to fail in a typical AbNET network, it is necessary for transmissions on all links of a node to fail. Suppose there is a hypothetical network in which a node has 3 links, each connected to a master 10 hops away. (The typical degree of connectivity for each AbNET node is three. An AbNET node is rarely connected to more than three other nodes, and an AbNET node is not needed if it connects only two other nodes.) The probability of one of the 10 links *in all three paths* dropping the packet is $(10\times10^{-6})^3$, or 10^{-15}, assuming there is no interconnection across paths. (Even this could be regarded as a worst-case scenario, as interconnection would improve the reliability.)

With flooding, it will be so long between dropped packets that other causes will predominate. The improvement due to the overconnected nature of the network is apparent, and justifies the non-use of error-detection-based retransmission at the data link layer.

While the numbers above are simple estimates based on assumed topologies, it is clear that the network is extremely robust if all the nodes are powered up. Power failure in part of the network is much more likely in the case of the power system application (where storm damage can black-out large areas), than in an industrial application, where the nodes are likely to be less remote. An industrial application can be expected to drop few packets.

6.6 PROTOTYPE SYSTEM

Licom, a company in Herndon, VA, produced a number of prototype circuit boards for AbNET. Each board had 4 pairs of fiber connectors, and 4 RS-232 connectors The firmware was programmed to execute the AbNET protocols and the Ping service. The board is shown below, in Figure 3-9.

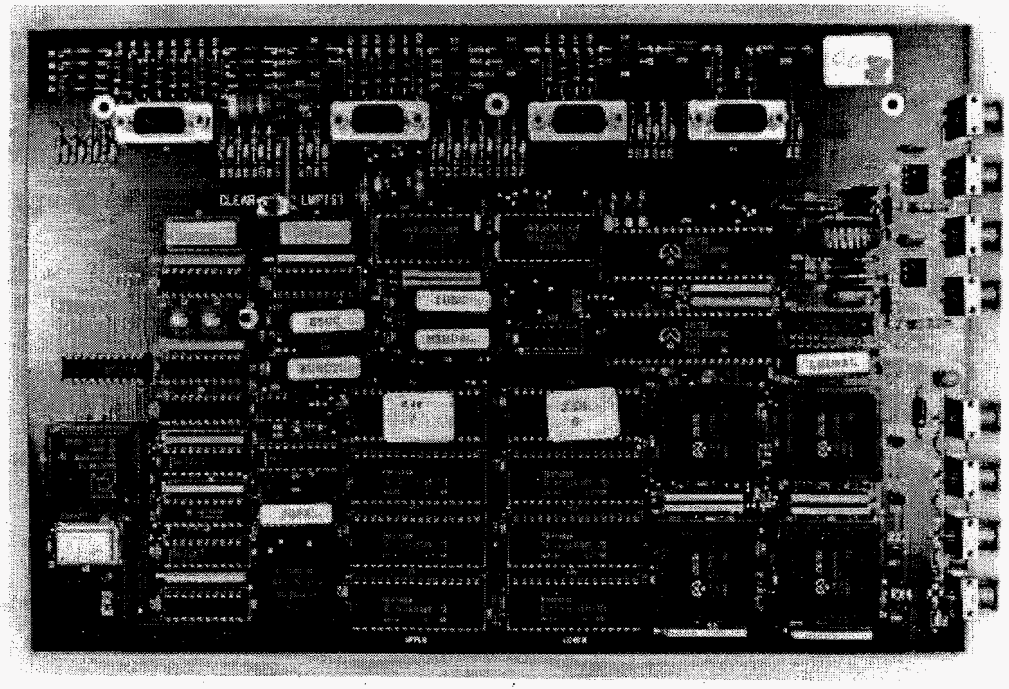


Figure 3-9. Licom's prototype AbNET board

The board is intended to be added to an existing RTU design, likely by installing it on the inside of the door. Power for the board is about 15 W, and it is designed to accept power over a wide range of input voltages from the RTU.

Two of the RS-232 connectors are for linking to RTUs. These are provided with complete surge suppression circuits in accordance with the relevant IEEE standards. The other two RS-232 sockets are for other uses. One is designated for monitoring the operation of the AbNET board itself, by utility personnel. The other is intended as a connection to the CEBus equipment that will be added for communication with the customer. This arrangement is shown in Figure 3-10.

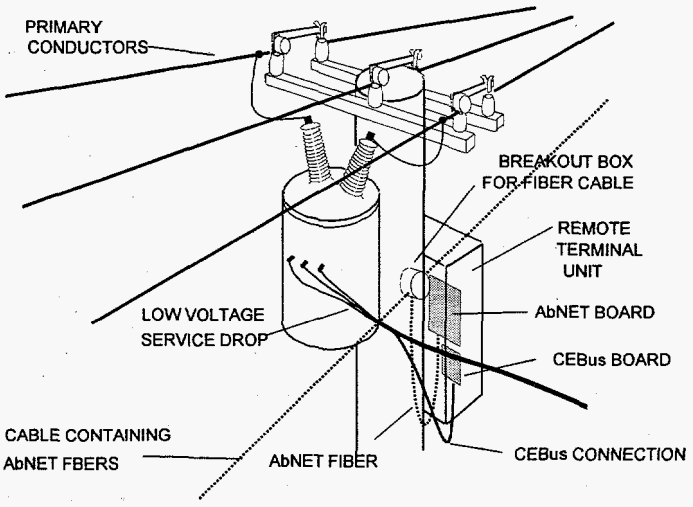


Figure 3-10. Possible arrangement of Licom's AbNET board in an RTU on a distribution pole, with CEBus board added

Fault tolerance is enhanced by error-detecting codes, added at 2 levels in the AbNET messages. First, the AbNET message itself includes error detection coding in the header. This packet is then embedded inside an SDLC packet that includes its own error detection codes. The message on the fiber side is as shown in Figure 3-11.

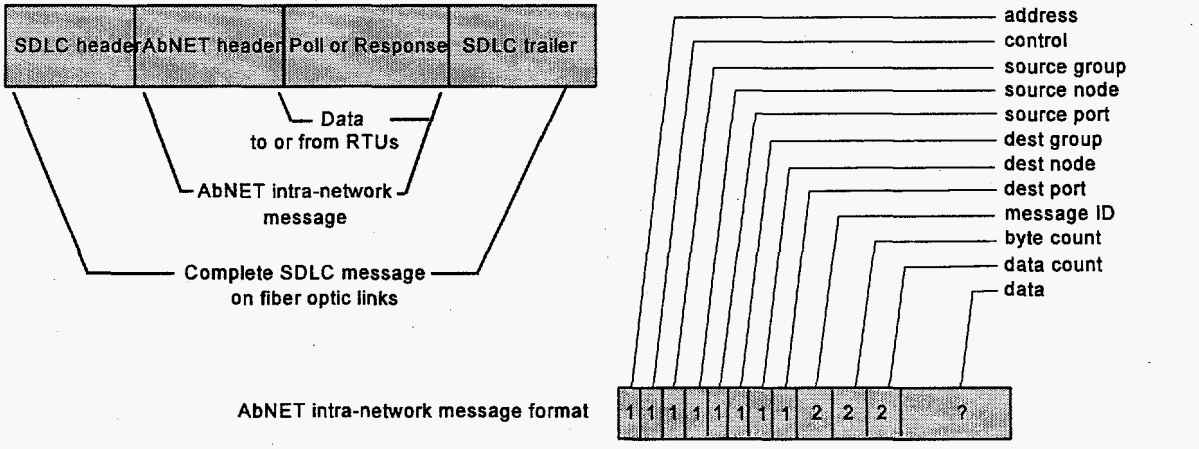


Figure 3-11. The AbNET message header contains group identification. It allows nodes to be selective about whose traffic they will handle. The AbNET package is embedded in a longer message, with its own header, for transmission on the fiber

The AbNET system has been tested in a lab-scale demonstration at both the John Dolan Laboratories of American Electric Power, and at the Jet Propulsion Laboratory. CEBus conversion has been implemented and tested.

Technically, the AbNET system is a success. So far, it has not been a commercial success. There are two possible reasons for this. First, the utilities have concluded that radio-based communications can be made less expensively than AbNET. This fact his a result of the development of UHF radio technology that did not exist when the AbNET work began. The utilities have not yet discovered that the same radio systems are too narrow in bandwidth to function in a polling scheme, and that they need report-by-exception to cope with the bandwidth limitations. These limitations of low-power UHF radio will not become evident until the attempt is made to make distribution automation work with more than one or two functions, or the service area covered is increased so that a larger number of nodes is included.

The other reason AbNET may not have been widely applied is that right about the time that AbNET was developed, the regulatory environment was changed. Deregulation is around the corner, and this has added uncertainty to the planning process. It has also meant that some of the common functions of distribution automation are now stretched across company boundaries. Load management, for example, which used to be viewed as a way a utility could defer peak load (and thereby defer additional generation capacity) operates to the benefit of the generating entity. Revenue is decreased for the distribution company. Successful load management would now require cooperation between the generating company and the distribution company, and the latter would likely expect compensation for lost revenue.

As the dust settles on the deregulation issue, it is possible that AbNET will come to be viewed as a reasonable solution to the communication problem. If companies do split horizontally, and concentrate on either generation or distribution (with the pool operator concentrating on transmission), then it may be that distribution will cease to be the poor relation in the electricity supply industry, and better management of assets will be needed.

6.7 CONCLUSIONS

The design of a fiber optic communication network for monitoring and control in power distribution systems has been discussed. By appropriate choice of protocols, a fault-tolerant system that operates in any arbitrary network configuration has been devised.

The network, called AbNET, is a packet-based distributed protocol system. Flooding is used for maximum failure tolerance. Hierarchical (master-slave) polling controls access to the system. This supports many data acquisition and control applications. The protocols allow multiple adjacent masters to share resources. A low-level network service reports the network's configuration to the master, where changes can be logged, and action taken if needed.

The system is transparent to the user, and maintains no record of the clients it serves. It is fast enough for all present and forseeable distribution automation applications, and many industrial control applications.

AbNET is intended to provide the communication infrastructure for any likely application, and will not require modification as more applications are added in the future. By appropriate choice of protocols, a fault-tolerant system that operates in any arbitrary network configuration has been devised. The protocols are matched to the requirements of distribution automation.

- The FLOODING algorithm guarantees
  - the maximum reliability and failure resistance—no other algorithm is better
  - the minimum transmission delay—the shortest route is always explored
- The NO-HANDSHAKE algorithm ensures that
  - communications lines and I/O processors are never busy with unpredictable communications because of link or node problems
- NO-HANDSHAKE and ANTIBODY algorithms together ensure
  - each link carries messages once and only once, predictably
- The POLLING approach ensures that
  - communications traffic is completely under the control of the central unit—no RTU can access the network (other than as a repeater) unless so instructed by the master station
- The HIGH SPEED allows
  - rapid sensing of the entire system
  - communication overhead and error detection codes without impairing system performance
  - communications to serve multiple functions
- The COMBINATION of fiber optics and PLC allows
  - low-cost implementation of low-performance applications such as load management
  - extension to new functions involving customer-owned equipment (resetting clocks etc)
- Services such as the PING mean that
  - the system configuration is automatically reported when AbNET is first powered up
  - unexpected changes in the system configuration are automatically recorded
  - adding new nodes or new lines is simply a matter of connecting them
  - Network performance can be readily monitored

The AbNET protocols were designed to control the flow of information in a network whose configuration was unusually complex, since it was determined by a pre-existing power network. To be useful in the power system, the solution had to be fault tolerant. As it happened, the design that resulted from following the power network led readily to a robust communication network.

VII. REFERENCES

The AbNET protocols are protected by U.S. Patents 5 341 372, 5 416 777, 5 434 865, 5 457 689

Adolfsson, M., Einvall, C.H., Lindberg, P., Samuelsson, J., Ahlgren, L. and Edlund, H., "EHV Series Capacitor Banks. A New Approach to Platform to Ground Signalling, Relay Protection and Supervision," Paper No. 88 SM 551-4, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

"Annual Industry Forecast," Special Report, Electrical World, October 1988.

Aucoin, M., Zeigler, J. and Russel, B.D., "Feeder Protection and Monitoring System, Part I: Design Implementation and Testing," "Part II: Staged Fault Test Demonstration," IEEE Trans. PAS, June 1985, pp. 1456-1462.

"Automatic Supervisory and Data Acquisition Systems for Electric Generation, Power Utilization, and Power Conversion Stations," described in an IEEE working paper, 1980.

Baran, M.E. and Wu, F.F., "Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing," Paper No. 88 SM 556-3, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Barnett, R.M., A Preliminary Estimate of Future Communications Traffic for the Electric Power System, JPL Publication 81-41, April 1981.

Bewley, L.V., Traveling Waves on Transmission Systems (2nd Ed.), Dover, New York, 1963, pp. 90-105.

Black, R.W., Formby, J.R. and Walker, R., "The Application of Feeder Automation to Existing and Future 11 kV Urban Networks," CIRED 1985: 8th International Conference on Electricity Distribution, May 20-24, 1985, Brighton, UK.

Bohlin, P., Edvinsson, M., Lindbergh, G. and Lundqvist, C.G., "Successful Implementation of a Nation-wide Load Management System," IEEE Transactions on Power Apparatus and Systems, November 1986, pp. 90-95.

Boice, D.G., Gursky, R.J. and Trad, J.B., "Cost of Electrical Power System Losses for Use in Economic Evaluations," Paper No. 88 SM 644-7, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Buch, J.F., Easley, J.H., Ezer, E., McCall, L.V., McLaughlin, P.K., Seebald, R.C., Stovall, J.P., Triplett, W.E. and Wolff, R.F., "Bibliography of Distribution Automation 1969-1982,"

IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 6, June 1984, pp. 1176-1182.

Bunch, J., Guidelines for Evaluating Distribution Automation, EPRI Publication No. EL-3728, Research Project 2021-1, Final Report, November 1984.

Bunch, J.B., Chen, A.C.M., Jenkins, D.R. and McCoy, J.L., "PROBE and Its Implications for Automated Distribution Systems," Proceedings of the American Power Conference, 1981.

Burke, J.J., "Cost/Benefit Analysis of Distribution Automation: Evaluation and Methodology," T&D Automation Conference and Exhibition, St. Louis, MO, September 27-29, 1988.

Caldwell, R.W., Adler, R.B., Alford, A., Bunch, J.B., Carr, J., Donaldson, D., Klein, K.W., McCall, L.V., Patton, J., Redmon J., Saccany, R. and Shelton, P. (1982), "The Distribution System of the Year 2000," IEEE Transactions on Power Apparatus and Systems, Vol. 101, August, pp. 2485-2490.

Carr, J., "Communication Requirements for Distribution Automation," Panel Session on Automated Distribution, IEEE Conference on Overhead and Underground Transmission and Distribution, Minneapolis, MN, September 20-25, 1981.

Crawford, M., "Back to the energy crisis," Science, Vol. 235, pp. 626-7, February 1987.

Critical Analysis of European Load Management Practices (1977), Energy Research and Development Administration, CONS/1168-1, available from NTIS.

DeLoach, B.C., Jr. and Gordon, E.I., "Optically Powered Implant in the Human Body," Western Electric Technical Digest, No. 64, pp. 17-18, October 1981.

DeLoach, B.C., Miller, R.C. and Kaufman, S., "Sound Alerter Powered Over an Optical Fiber," Bell System Technical Journal, No. 57, pp. 3309-16, November 1978.

Distribution Automation, IEEE Working Group on Distribution Automation (Eds.), IEEE Tutorial Course text No. 88EH0280-8-PWR, 1988.

Dugan, R.C., Thomas, S.A. and Rizy, D.T., "Integrating Dispersed Storage and Generation With an Automated Distribution System," IEEE Trans. PAS, Vol. 103, No. 6, June 1984, pp. 1142-1146.

Fiske, G.S., Law, E.T.K. and Seeto, D.Q., "The Economic Analysis of Load Management: The Case of Cycling Residential Air Conditioners," IEEE Trans. PAS, December 1981, pp. 4725-4732.

Gardner, F.M., Phaselock Techniques, New York, Wiley & Sons (1966).

Gilchriest, C., RF Model of the Distribution System as a Communication Channel, Final Report, Contract No. 955647, July 28, 1982. Performed by General Electric for the Jet Propulsion Laboratory.

Groghan, J.F., Jenkins, D.R., Rushden, F.A., Bunch, J.B. and Gurr, G.P., "PROBE-A Demonstration of Substation and Distribution Automation," Proceedings of the American Power Conference, 1977, Vol. 39, pp. 1279-1287.

Hagmann, W., "A Spread Spectrum Communication System for Load Management and Distribution Automation," Paper No. 88 WM 060-6, presented at the IEEE PES Winter Meeting in New York, NY, January 31 - February 5, 1988.

Hall, P., "An Optically Powered Sensor Network," IEE Colloquium on Distributed Optical Fiber Sensors (Digest No. 74), 1986.

Horstein, M. and Barnett, R., Satellite Applications to Electric-Utility Communications Needs, JPL Publication 81-104, December 1, 1981.

Johnson, B.L., Kissinger, D.W., Koepfinger, J.L., Stadlin, W.O. and Masiello, R.D., "Automated Distribution Network Control," Power Industry Computer Applications, pp. 276-283, 1975.

Johnston, A.R., Jackson, S.P., Kirkham, H. and Yeh, C., Power System Applications of Fiber Optic Sensors, JPL Publication 86-22, June 1986, Jet Propulsion Laboratory, Pasadena, California.

Kendrew, T. J., and Marks J. A., "Automated Distribution Comes of Age", IEEE Computer Applications in Power, Vol. 2, No. 1, January 1989, pp. 7-10.

Kirkham, H. and Johnston, A.R., Electric and Magnetic Field Meters Developed for the U.S. Department of Energy, JPL Publication 88-1, 1988, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R., Jackson, S.P. and Sheu, K., AC and DC Electric Field Meters Developed for the U.S. Department of Energy, JPL Publication 87-20, 1987, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R., Lutes, G., Daud, T. and Hyland S., Power System Application of Fiber Optics, JPL Publication 84-28, Jet Propulsion Laboratory, Pasadena, California, 1984.

Lee, R.E. and Brooks, C.L., "A Method and Its Application to Evaluate Automated Distribution Control," IEEE Transactions on Power Delivery, Vol. 3, No. 3, July 1988, pp. 1232-1240.

Lewis, S.M., "Distribution Automation Developments," T&D Magazine, Feb. 1981, pp. 24-25.

Liu, C. C., Lee, S.J. and Vu, K., "Loss Minimization of Distribution Feeders: Optimality and Algorithms," Paper No. 88 SM 580-3, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Lyons, P.C. and Thomas, S.A., "Microprocessor-based Control of Distribution Systems," IEEE Trans. PAS, December 1981, pp. 4893-4900.

Ma, F.S., Isaksen, L. and Kuliasha, M., "Impact of Dispersed Supply Management on Electric Distribution Planning and Operations," IEEE Trans. PAS, Vol. PAS-98, No. 5, Sept./Oct. 1979, pp. 1531-1539.

Markell, L.C., Summary of the Athens Automation and Control Experiment (Contract No. DE-AC01-84CE76248), Washington, D.C., Department of Energy, Office of Energy Storage and Distribution, 1987.

McCall, L.V., "A Distribution Automation Demonstration Project," IEEE Transactions on Power Apparatus and Systems, Vol. 100, No. 4, pp. 1744-1749, April 1981.

McCall, L.V. and Chambers, B.J., "Scarborough Distribution Automation Project: Implementation and Preliminary Performance Experience," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-104, No. 10, pp. 2759-2763, October, 1985.

McGlade, S.M. and Jones, G.R., "An Optically Powered Vibrating Quartz Force Sensor," G.E.C. Journal of Research (U.K.), Vol. 2, No. 2, pp. 135-8, 1984.

Monteen, L.D., Survey of Communication Alternatives for Distribution Automation and Control Systems, University of Tennessee, April 1979.

Monteen, L.D., Lawler, J.S., Patton, J.B. and Rizy, D.T., "Impact of Automation on the Reliability of the Athens Utilities Board's Distribution System," Paper No. 88 WM 087-9, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Morgan, M.G. and Talukdar, S.N., "Electric Power Load Management: Some Technical, Economic, Regulatory and Social Issues," Proceedings of the IEEE, Vol. 67, No. 2, February 1979, pp. 241-313.

Nelson, W.T., "Operation of a Successful Load Management System," IEEE PES Paper No. 81 SM 401-9, 1981.

Nightingale, D. and Satinski, S., Analysis of IEDMS Communication System, JPL Report D-1697 (internal document), June 28, 1984.

Ohte, A., Akiyama, K. and Ohno, I., "Optically-powered Transducer With Optical Fiber Data Link," SPIE, Vol. 478, Fiber Optic and Laser Sensors II, 1984, pp. 33-38.

Platts, J., "Electrical Load Management: The British Experience," IEEE Spectrum, February 1979, pp. 39-42.

Purucker, S.L., Reddoch, T.W., Detwiler, J.S. and Monteen, L.D., "The Design of an Integrated Distribution Control System," IEEE Trans. PAS, Vol. 104, No. 3, March 1985, pp. 745-752.

Rau, N.S. and Graham, R.W., "Analysis and Simulation of the Effects of Controlled Water Heaters in a Winter Peaking System," IEEE Trans. PAS, Vol. PAS-98, No. 2, March/April 1979, pp. 458-464.

Reed, J.H., Broadwater, R.P. and Chandrasekaran, A., "Load Control Experiments With Heat Pumps During the Winter," Paper No. 88 WM 094-5, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Reed, J.H., Nelson, W.R., Wetherington, G.R. and Broadaway, E.R., "Monitoring Load Control at the Feeder Level Using High Speed Monitoring Equipment," Paper No. 88 WM 095-2, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Reed, J.H., Thompson, J.C., Broadwater, R.P. and Chandrasekaran, A., "Analysis of Water Heater Data from Athens Load Control Experiment," Paper No. 88 SM 596-9, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Rhyne, V.T., "Field Demonstrations of Communication Systems for Distribution Automation," DOE/NBB-0012, June 1982.

Rizy, D.T., Lawler, J.S., Patton, J.B. and Nelson, W.R., "Measuring and Analyzing the Impact of Voltage and Capacitor Control With High Speed Data Acquisition," Paper No. 88 WM 098-6, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Rogowski, V.W. and Steinhaus, W., "Die Messung der Magnetischen Spannung," Archiv für Elektrotechnik, Vol. 1, No. 4, pp. 141-50, 1912 (in German).

Russell, B.D. (Ed.), "Communication Alternatives for Distribution Metering and Load Management," Record of Panel Presentations, IEEE PES 1979 Summer Meeting, prepared by Application of New Technologies in Substation Control Working Group, Automatic and Supervisory Systems Subcommittee, Substations Committee. IEEE Transactions on Power

Apparatus and Systems, Vol. PAS-99, No. 4, July/August 1980, pp. 1448-1455.

Sakaguchi, T. and Matsumoto, K., "Development of a Knowledge Based System for Power System Restoration," IEEE Trans. PAS, Vol. 102, No. 2, February 1983, pp. 320-329.

Schweizer, P., Neveux, L. and Ostowsky, D.B., "Optical Fiber Powered Pressure Sensor," SPIE, Vol. 798, Fiber Optic and Laser Sensors II, 1987, pp. 82-85.

Shirmohammadi, D. and Hong, H.W., "Reconfiguration of Electric Distribution Networks for Resistive Line Loss Reduction," Paper No. 88 SM 598-5, presented at the IEEE PES Summer Meeting in Portland, OR, July 24-29, 1988.

Survey of Utility Load Management and Energy Conservation Projects (1977), U.S. Department of Energy ORNL/Sub-77/13509/2, Available from NTIS.

"Tax Gas", The Economist, December 24, 1988, pp. 13-14.

U.S. Energy '88: Countdown To The Next Crisis, United States Energy Association, Washington, D.C., April 1988.

Vaisnys, A., A Study of a Space Communication System for the Control and Monitoring of the Electric Distribution System, JPL Publication 80-48, May 1980.

Weers, D.D. and Shamsedin, M.A., "Testing a New Direct Load Control Power Line Communication System," IEEE Trans. Power Delivery, Vol. PWRD-2, No. 3, July 1987, pp. 657-660.

Witting, H.L., Development of a Cesium-vapor-lamp System for Triggering Photo-thyristors for Power Conversion Applications in Electric Utility Substations, Final Report, EPRI Publication No. EL-3140, June 1983.

Section 4: NEW ANALOG DATA LINKS

SUMMARY

New optical data links that could be optically powered are described in this section of the Final Report. There are two versions: each has good frequency response and minimal filtering requirements. In one, a conventional FM pulse train is used at the transmitter, and a novel form of phase locked loop is used as demodulator. In the other, the FM transmitter is replaced with a pulse generator arranged so that the period between pulses represents the modulating signal. Transmitter and receiver designs, including temperature compensation methods, are presented. Experimental results are given.

4 AN IMPROVED FM LINK

4.1 INTRODUCTION

In 1969, when the author of this report was working on the stability of ac power systems with embedded dc lines, the need arose to measure the power angle of a generator. Power angle is the term used to describe the phase angle between the terminal voltage of the machine and the voltage behind synchronous reactance. The second of these voltages is not observable; it is the voltage induced in the stator winding. As a surrogate of the power angle, the group of which the author was then part had been using the derivative of the rotor speed.

The derivative of rotor speed had shown itself to be an adequate surrogate. However, this measurement was not easy. The derivative operation is inherently noisy. The circuit design became one in which the need for a rising frequency characteristic (to obtain the derivative) had to be balanced by the need for a falling characteristic (to keep noise at an acceptable level). To reduce the high-frequency component of the signal entering the derivative circuit, a simple ramp-sample-hold (RSH) arrangement was used. Figure 4-1 shows the idea.

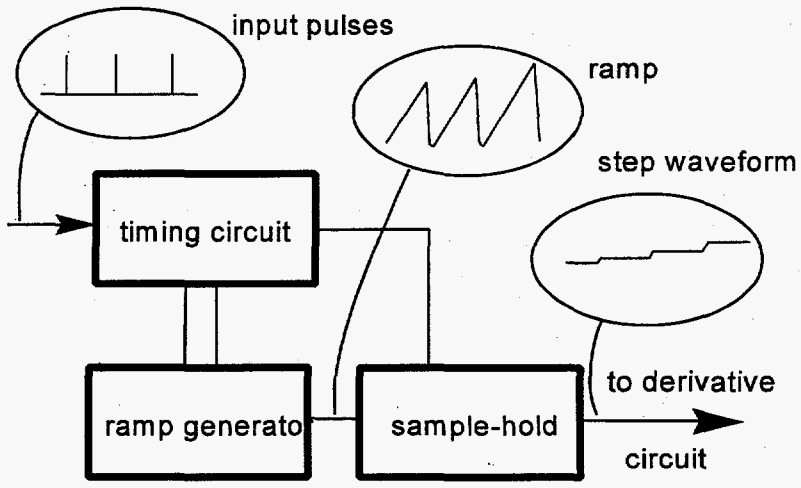


Figure 4-1. Block diagram of angle derivative circuit

The train of input pulses cause timing circuit to instruct the ramp generator to reset and re-ramp. (In practice, this sequence requires two operations. The distinction is unimportant in terms of understanding the operation.) If the machine is slowing down, as shown here, the input pulses get further apart, and the ramp has time to grow to a larger value. Just before the ramp resets, the

sample-hold circuit is instructed to acquire the new ramp value. The result is that if the speed is constant, the output of the circuit is a steady dc level, but if the speed is changing, a step waveform is generated. If the machine is "swinging," or oscillating either side of a steady value, the output is a stepped sine-wave that certainly seemed to resemble the curve of machine speed.

The filtered step waveform is a relatively "smooth" function to apply to the derivative circuit, and it simplified the filtering problem. This kind of system was first used to demonstrate that the stability of an ac system can be improved by controlling the part of the power flow on an embedded dc line in proportion to the derivative of the power angle (Dougherty and Kirkham, 1970).

In fact, the resemblance between the angle measuring system output and the relative speed is somewhat illusory. The amplitude of the ramp is proportional to the time between the set and reset pulses, essentially *inversely proportional* to the speed. This signal can be electronically inverted, but that only means multiplying by −1. The system works only because in this application the speed of the machine varies very little about its nominal value. The reciprocal relationship between the speed and the electronic representation of it is unimportant if the reciprocal curve is viewed as piecewise linear. Figure 4-2 shows this.

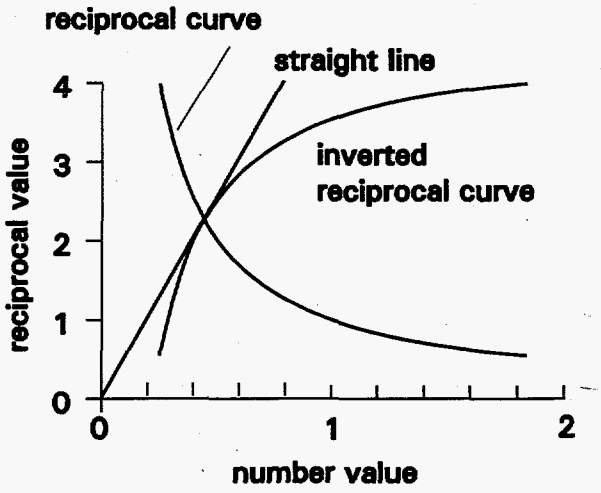


Figure 4-2. Approximation of proportionality by the inverted reciprocal

Now we move closer to the present time. For most of the early development of the optically powered electronic measuring systems, we had used FM to transmit the data from the measuring system to the receiver. At most "ordinary" frequencies, a phase locked loop can be used as a demodulator of FM signals, and a dynamic range of about 60 or 70 dB can readily be obtained. However, to conserve power, we had used a very low carrier frequency, and this had limited the bandwidth of the system.

Nominally, the process of demodulation is the extraction of the modulation information from the

modulated carrier. Whatever demodulation process is used, a certain amount of the carrier is bound to remain. When the carrier frequency is decreased, unless the bandwidth is also decreased, the problem of separating this residual carrier from the desired output becomes more difficult.[1]

Normally, the separation process is performed by frequency-selective filters in the analog domain. While this approach is practicable, the drawbacks of complex filtering and modified phase response make it worthwhile to investigate other solutions.

Consider the basic phase locked loop (PLL). A phase locked loop consists of 3 essential elements[2]: a phase detector or comparator, which compares the phase of an incoming signal with that of a locally generated signal; a loop filter, which smooths the phase comparator output to produce an error voltage; and a voltage controlled oscillator, whose frequency is a function of the applied error voltage. They are typically arranged as in Figure 4-3.

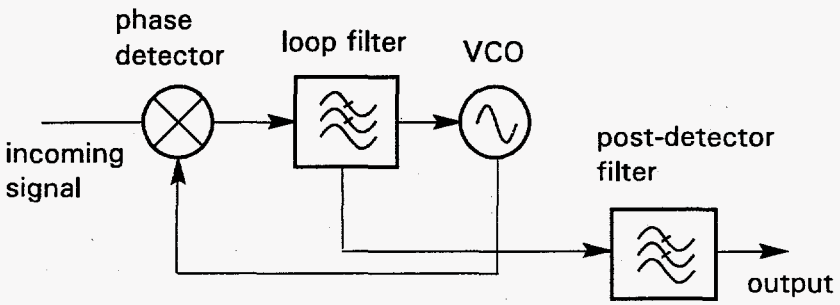


Figure 4-3. Typical phase locked loop

[1] An analogous situation arises in reconstructing analog waveforms from the information on a digital Compact Disc. An elegant solution in many modern CD players is to use the technique of oversampling. This can be explained as follows. Imagine the output waveform to be reconstructed by means of a sample-hold system, or staircase generator. Since the sample rate on CDs is about 44 kHz, the problem is to remove the remnant 44 kHz component without affecting the desired audio at, say, 20 kHz. By interpolating between adjacent values of the signal, additional steps can be created in the reconstructed output. The apparent sampling rate can be increased by a factor of 2, 4 or even 8 times. This means that the undesired residual can be placed as much as 16 times higher in frequency than the desired output. A very simple filter can then accomplish the task of separating such signals.

This is not an option open to us, since we are dealing with real-time signals, not stored data. The delay involved in waiting for samples for interpolation is unacceptable.

[2] There is also the possibility of non-essential elements, such as frequency dividers (which can be used in the forward or the feedback parts of the loop), and presettable counters, which can be similarly applied. These components do not affect the present discussion.

The phase comparator and loop filter determine the lock range, the capture range and the noise performance of the loop. If the incoming signal has a large signal-to-noise ratio, one is less concerned about noise performance. When used as an FM demodulator, the loop filter has to be a compromise between flat frequency response and rejection of the unwanted phase comparator products. As the spectral frequency of the incoming modulation approaches the carrier frequency, this compromise becomes difficult to resolve.

Note that with the typical multiplying-type phase detector, the lowest frequency component of the phase detector output is at twice the carrier frequency, f_0. However, the amplitude of this component is large, regardless of the amplitude of the modulating signal, because this kind of phase detector switches between two extreme states. Typically, the signal applied to the low-pass loop filter is a square wave with peak amplitude equal to the local power supply voltage.

Stability considerations dictate the use of simple low-pass or lag-lead loop filters. The filter of a wideband loop removes little of the energy of the signal at $2f_0$. For an undistorted output, a post-detector filter of considerable quality is required. Such a filter was shown in Figure 4-3.

The first of our two revisions to the optical link greatly reduces the filter requirements. By replacing the phase detector and loop filter with a sample-hold system, as in Figure 4-4, the ripple component of the phase detector output is greatly reduced. With this modified loop, the total ripple energy is dependent on the signal amplitude, and is largely in the high harmonics. The idea can be traced directly to the power angle problem back in 1970, though at that time the application did not involve a PLL.

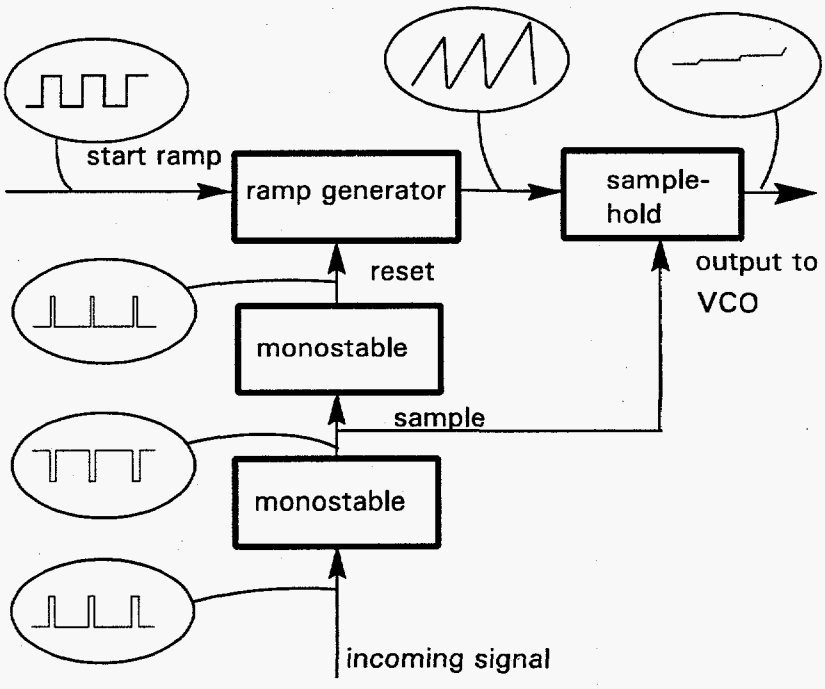


Figure 4-4. Block diagram of sample-hold system

Operation is as follows. Suppose that a reset pulse has just set the ramp generator to zero. The system is static until a signal comes from the VCO. This starts the ramp generator, which can be a simple op-amp integrator. A short time later, a pulse comes from the incoming signal line. This triggers a monostable oscillator, which produces a short pulse that tells the sample-hold circuit to hold. The voltage that appears on the output of the sample-hold is thus a measure of the time difference between the VCO and the incoming signal. The trailing edge of the monostable triggers a second monostable, arranged to reset the ramp generator. The second monostable serves to delay the reset until the sample-hold has acquired a new value. The sequence repeats.

The timing diagram (Figure 4-5) shows the relationship between the various signals. Starting at the bottom of the diagram, the incoming signal causes the first monostable to generate the SAMPLE pulse. This puts the current value of the ramp generator on the output of the sample-hold circuit. The trailing edge of the SAMPLE pulse starts the RESET pulse, which causes the ramp generator to discharge. The ramp circuit starts to charge in response to the VCO signal. For a decreasing frequency of the incoming signal, the output is the familiar staircase waveform.

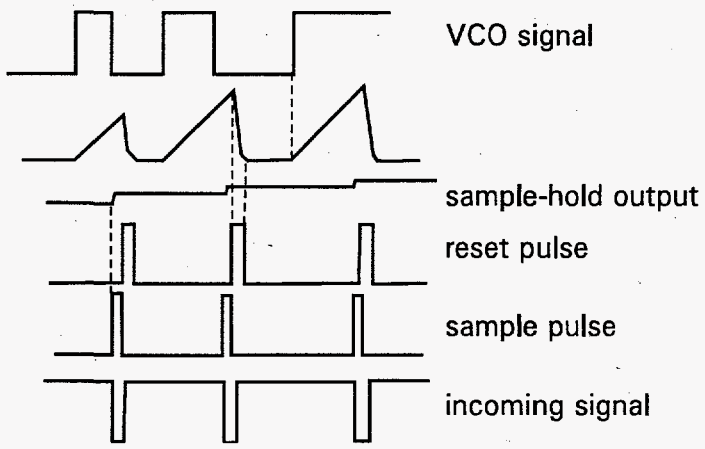


Figure 4-5. Timing diagram

Because the circuit is inside a phase locked loop, the time difference between the incoming signal and the VCO is a measure of the phase difference. The ramp and sample-hold circuits together perform the function of phase detector.

4.1.1 Higher order loop

It is possible to look at this another way. The interval between the pulses of incoming data in a frequency modulated data link contains information about what is going on at the transmitter. The interval between any two adjacent pulses can therefore, in principle, be used to estimate the state of the transmitter. Some of the techniques we have used to implement the new demodulator system are borrowed from the world of state estimation.

In essence, the device attempts to duplicate in the demodulator the events going on in the modulator. The interval between pulses in the telemetry link is used as an estimate of the voltage on the control terminal of the VCO. A voltage derived from this estimate is applied to the demodulator VCO, which therefore mimics the transmitter VCO. The accuracy of reproduction is controlled by the fact that the loop is closed: any deviation represents an error that will generate a correcting signal in the phase locked loop.

The simplest way of estimating what is going on in the transmitter is to assume that the control voltage applied to the VCO did not change during the period between output pulses. This is what was described above; it results in a loop that approximates the performance of a conventional first order loop, although the estimate of the modulating signal is known as a zero order estimate. (From now on, we must distinguish between the order of the estimator and the order of the loop.) If the modulating signal is constant there is no error. If the signal is changing at a constant rate, there is an error due to the step-like nature of the approximation. This can be seen in Figure 2-4.

If information about the way the period changes from interval to interval is incorporated, a first order estimate results. In a first order estimator, the assumption is made that the change during the period is constant. The error is zero when either the signal or its rate of change is constant. The approximation consists of small ramp segments, rather than small constant segments. The design of a loop using such an estimator is discussed next.

4.1.2 First order estimator: loop implementation

In the general state estimation scheme, the first order estimate can be obtained by incorporating information about the difference between adjacent samples of the data. For example, suppose that between one sample and the next, the value is found to have increased by 10%. The zero order system would assume that the last measured sample is the best estimate of the next. The output of the system is constant until the next sample arrives. The first order estimator would assume instead that there will be a further 10% increase in value during the next interval. The output is a ramp, starting at a value that represents the last measurement, and calculated to increase the output by 10% during the next period. The hardware for this is clearly going to be a little more complex than for the zero order system.

Our first order estimator functions as follows. Control signals (short pulses) are derived, as before, from the incoming signal by means of simple delay circuits. The details of this are not shown in the block diagram, Figure 4-6. A zero order system is implemented, as before, by means of a ramp generator and a sample/hold circuit. The ramp generator of this part of the system now contains circuitry to enable it to hold the output constant for a brief interval before it is reset.

To make the estimator first order, additional hardware computes the difference between adjacent

intervals. An analog difference circuit continuously monitors the difference between the output of the zero order hardware just described and the input to its sample/hold circuit, see Figure 4-6. At the instant that the zero order sample/hold circuit is instructed to sample, this difference represents the change in the signal between two adjacent intervals. This value is therefore sampled just before the sample in the zero order part of the system.

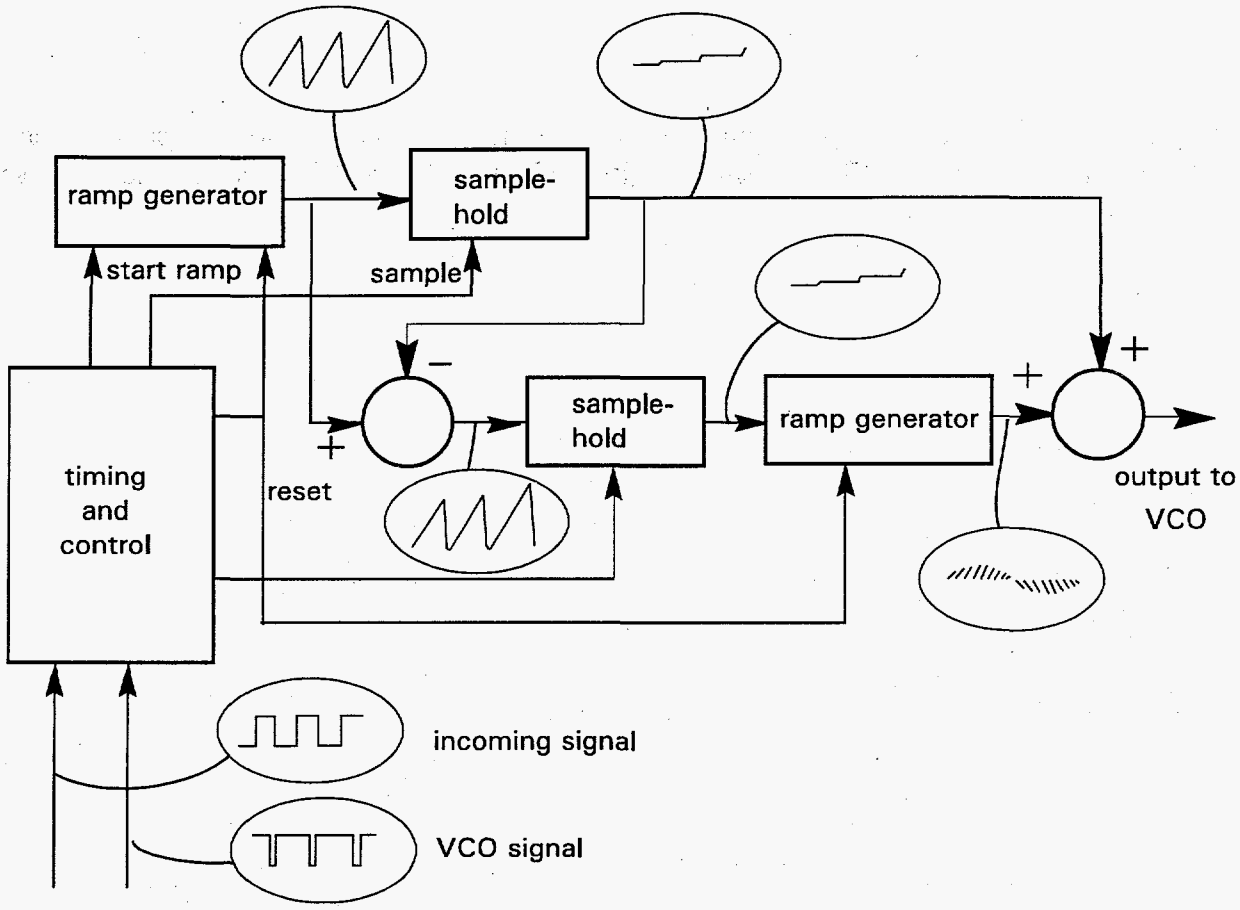


Figure 4-6. First order estimator loop

A second ramp circuit is used to generate a ramp whose slope depends on the output of the difference sample/hold circuit. The direction of the ramp depends on the sign of the difference, and the steepness of the ramp is controlled by the magnitude of the difference signal. These waveforms are shown in Figure 4-7 for a high modulating frequency (few samples per cycle).

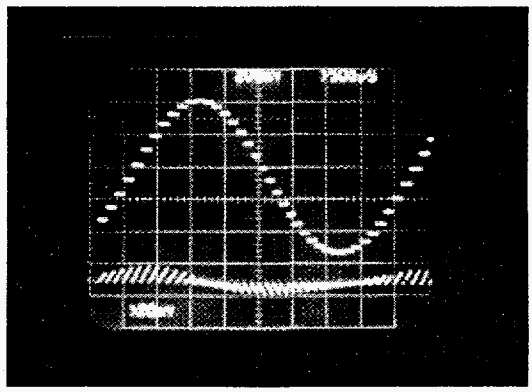


Figure 4-7. First order estimator waveforms

The first order estimate is obtained by adding the zero order estimate and the difference-generated ramp, ie by adding the two signals shown in Figure 4-7..

4.2 Zero order estimator loop - measured performance

The simplest implementation of our loop uses the zero order estimator, ie one that has no derivative compensation. This version has, nevertheless, significant performance advantages over conventional loops. To demonstrate this, we arranged loops of various kind, conventional as well as sample/hold, as demodulators of a low-frequency FM carrier. All loops were fed the same signal. The results stress the low noise content of the demodulated signal from our loop.

There are normally considered to be 2 kinds of phase locked loop, distinguished by their phase detector. A loop using the original multiplying-type phase detector is known as a Type I loop. This phase detector has the characteristic that, with lock at the VCO free-running frequency, the output waveform has unity mark:space ratio. After filtering, this signal averages to zero, satisfying the condition for lock at the free-running frequency. There is a 90° phase angle between the input signal and the VCO. At other frequencies in the lock range, there is a different phase shift between the input signal and the VCO signal. A non-zero error voltage is generated because the mark:space ratio changes.

In most implementations, the VCO waveform is a square wave. The phase detector output waveform depends on the signal waveform. If the signal is sinusoidal, the filtered phase detector output is a sinusoidal function of the phase angle. If the input signal is a square wave, the filtered phse detector output is a linear function of the phase displacement. These characteristics are well known. (See, for example Gardner, 1966.)

What is sometimes overlooked is that the unfiltered phase detector output is a large amplitude

signal. In applications where the input signal is amplitude limited before the phase detector, the output waveform is a square wave whose instantaneous amplitude is, for all practical purposes, equal to the system power supply voltage.

The other conventional phase detector (Type II) is based on the detection of zero crossings in the input and VCO signals. When locked at the free-running frequency, this kind of loop has 0° phase angle between the signal and the VCO. Their zero crossings are coincident. The phase detector output is zero. If the VCO should drift slightly, the zero crossing of one signal will move ahead of the other. Depending on the relative phase, the phase detector produces an output impulse of positive or negative polarity. The filtered pulse acts so as to correct the VCO frequency. If the drift (or frequency shift) is slight, these pulses are produced only occasionally. If the loop is operating further from the free-running frequency, these pulses may be produced every cycle.

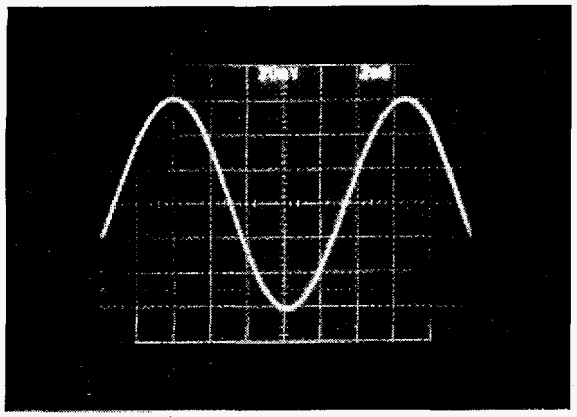
By extension of this naming convention, we designate our phase locked loop a Type III loop in the following discussion.

Figures 2-7 through 2-9 compare the output of the three loops for an input modulation of 20 mV. (This is about in the center of the usable range of input signals.) In Figure 2-7(a) the loop error voltage (ie, the output of the loop when used as a demodulator) for a Type I loop is shown. In operation, this signal would be filtered to remove the modulation content, in this case at 80 Hz. Figure 2-7(a) shows that the loop output is predominantly a square wave at the carrier frequency (about 7000 Hz), effectively pulse width modulated at 80 Hz. The signal spectrum is shown in Figure 2-7(b). Here it can be seen that the carrier is over 30 dB higher than the signal that is desired. This means that the filtering that is applied must cope with a "noise" over 30 times higher than the desired "signal".

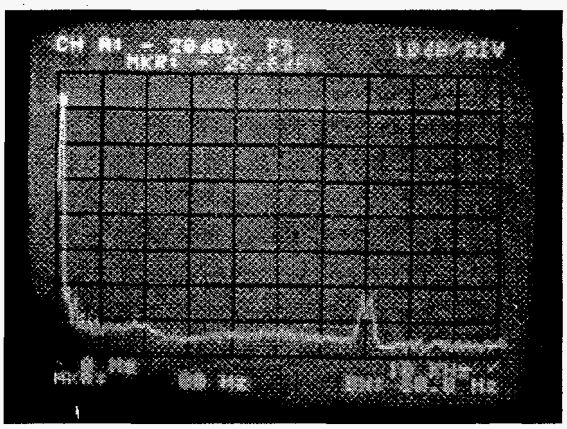
< >deletion

4.3 First order estimator - measured loop performance

As may be expected, the output of the loop with a first order estimator "looks" a lot better than the output of the zero order version. The discontinuities in the waveform are much smaller because of the ramp-like nature of the estimate. On the spectrum analyzer, some of the expected improvement disappears. There are two reasons for this. First, the first order estimator requires more digital switching signals, and these fast, short pulses tend to feed through into some of the analog circuitry. Second, the difference signal is a small value obtained by subtracting one large signal from another. Such a process is inherently noisy. Nevertheless, the carrier rejection properties of the loop are measurably better than those obtained with the zero order estimator. Figures 2-10 through 2-12 compare loops made with zero order and first order estimators.

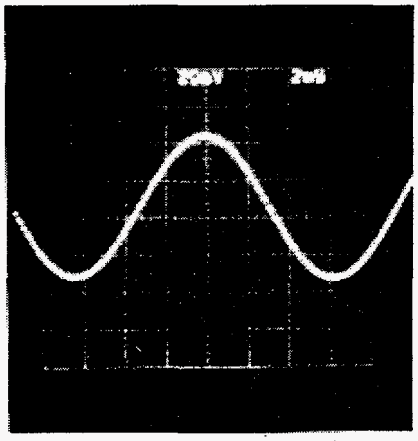


(a) Loop output signal

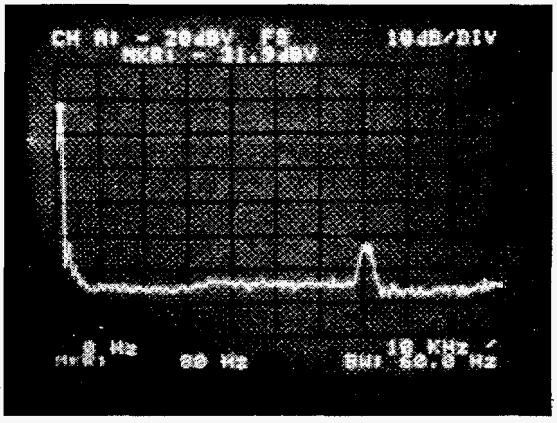


(b) Output signal spectrum

Figure 4-8 (a) and (b). Zero order estimator, V_{in} = 20 mV



(c) Loop output signal



(d) Output signal spectrum

Figure 4-8 (c) and (d). First order estimator, V_{in} = 20 mV

There is no doubt that the output of the loop is a remarkable reconstruction of the modulation. It should be remembered that the waveforms shown here are not filtered at all. Not only is the additional circuitry of an analog filter not required, but the unavoidable effects of such a filter on the frequency response and the phase response of the loop are avoided. These topics are discussed further below.

4.4 Additional observations

The phase locked loop implemented here contains both linear and nonlinear elements. Analysis of the loop performance is therefore somewhat difficult. However, its behavior can be considered by analogy with a conventional PLL, with some interesting conclusions in the area of stability and frequency response.

< >deletion

The frequency response of a conventional PLL is determined by the loop filter, which in general has low pass characteristics. At best, the design of the loop filter is a compromise between obtaining a wide enough bandwidth that the loop will follow the modulation and a narrow enough filter that the carrier itself is somewhat rejected. The order of a conventional loop is determined by the number of poles in its response, particularly in its filter. A first order loop (ie one pole) is obtained with no filter at all. (But the loop response contains a pole because of the integral relationship between the frequency and the phase of the VCO.) If a simple low-pass loop filter is used, resulting in a second order loop, and if the filter cutoff is set at a low frequency, the loop response will be narrowband. The loop will be unable to track a rapidly changing input. Further, as loop gain is increased, the response is underdamped, and the transient response is poor.

Lead compensation is normally used to solve this problem. A filter with a lag-lead characteristic results in a second order loop with improved transient response and stability. However, the carrier rejection is degraded, and additional filtering may be required outside the loop.

Our nonlinear system is slightly different. With the zero order estimator used in the phase detector, the loop response appears to be intermediate between first and second order. At low modulating frequencies, the phase detector is effectively instantaneous, and since there is no filter, the loop response is apparently first order. At high modulating frequencies, the time lag in the steps of the phase detector output make the response approximate a second order loop. As the loop gain is increased, there is a decrease in the damping. In fact, to some extent the loop gain can be used to tailor the frequency response. A discussion of feedback systems with zero-order hold is given by Elgerd (1967) and by Kuo (1967).

With our Type III system, analog filtering need not be used, although the possibility is not ruled out. Another possibility is the use of a stepwise signal for lead compensation derived in the same way as the signal for the difference ramp generator in the first order estimator. This could be applied in a summing circuit to tailor the loop frequency response. It is likely that the loop response would resemble that of an analog second order loop with lead compensation. We have made no attempt to investigate such a loop in our prototypes, so far. We note, however, that there is a measure of lead compensation inherent in the first order estimator.

4.5 Performance evaluation

While the waveforms and spectra presented above are indicative of improved performance, the usual way of assessing any improvement is in terms of the frequency response (or the step response) and the dynamic range of the link. In this section, we examine those aspects of the performance of a link using the new phase locked loop. In these examples, the carrier frequency is about 5 kHz.

4.5.1 Frequency and step response

The high frequency response can be adjusted between a peak and a roll-off by means of the loop gain. The step response, shown in Figure 2-13b, shows both overshoot and undershoot, perhaps evidence of non-linear behavior. The low frequency response is excellent, extending (as with conventional FM) down to dc. The high frequency response was found to be slightly amplitude dependent. It could be peaked, in the region of 1 kHz, by changing the loop gain.

4.5.2 Dynamic range

The dynamic range is shown in Figure 4-9, which was plotted at a frequency of 80 Hz, in the linear range of operation, and close to the 60-Hz that is of interest in most power system measurements. The transfer function of the zero-order loop with a simple two-pole low pass filter is within $<10\%$ from about 0.5 V down to a level of 200mV as shown.

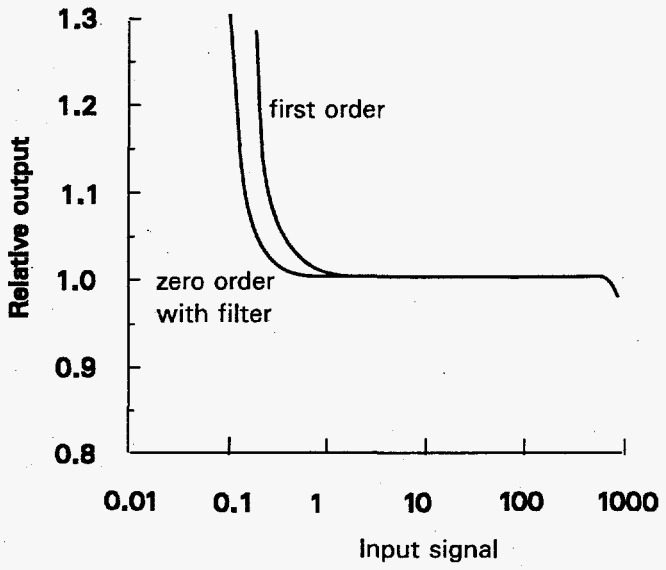


Figure 4-9. Dynamic range

The system has a dynamic range in the order of 70 dB, enough to make it useful in a number of measurement applications, but also enough to make characterization difficult. During the course of these measurements, it was necessary at various times to add attenuators (at the input) and an amplifier (at the output) and to change range on our measuring voltmeter.

4.5.3 Lock range

The input-frequency/output-voltage relationship of a phase locked loop is typically a sawtooth. As the frequency is increased slowly from an unlocked low frequency, the loop provides no output. The average output of the phase detector, filtered in the normal PLL, is zero. At some point, the loop will lock, and the output suddenly jumps to a large value. This is the value needed to pull the VCO to the input frequency. As the frequency continues to increase, the output voltage will track the frequency (providing a way to measure the linearity of the PLL) until, at some point, the loop loses lock, and the output voltage returns to zero. If the input frequency is now slowly decreased, the loop will reacquire lock, usually at a lower frequency than the frequency at which it lost lock. Lock will be maintained, usually to a lower frequency than the frequency at which lock was first acquired.

The foregoing description applies to a loop with a Type I phase detector. The slope of the characteristic is the reciprocal of the VCO gain. The type II detector will give the loop similar performance, but will often have a wider lock range, since the Type II phase detector is really a phase/frequency detector, and can apply a non-zero voltage to the VCO even when the loop is not locked. The test was performed on the type III loop, with the results shown in Figure 4-10.

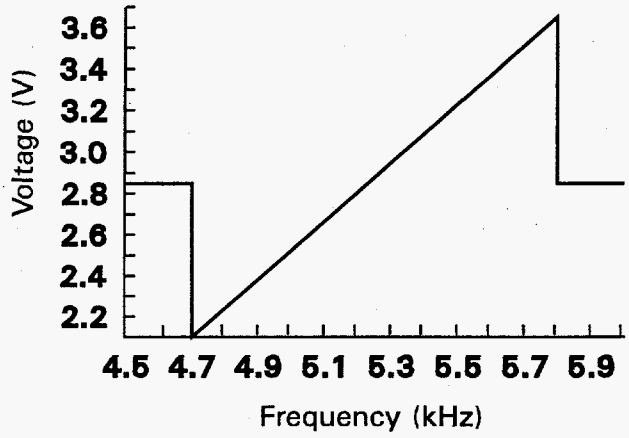


Figure 4-10. Lock range

The range for lock and acquisition are indistinguishable. The linearity is excellent, in fact the nonlinearity is observed to be $<0.3\%$, the uncertainty of our measurements. The lock range is large enough, in terms of the input frequency, to make the device useful as a tachometer.

4.6 Concluding Remarks

Detectors for FM are not new. Indeed, there are a large number of ways that an FM signal can be modulated and demodulated, the choice being dependent on the application. For use in commercial VHF radio, the "discriminator" was once the detector method of choice. Recently, the phase locked loop has been used in this application. Since the required output is band-limited at less than 20 kHz, it is not difficult to remove the PLL phase detector products, which are all above the IF of 10.7 MHz. As with conventional discriminators, a simple passive single-pole filter will perform well.

The constraint in our application is primarily at the modulator. The uncertainties in the power margin of the optical data link lead to the choice of frequency modulation, since the output is independent of the input signal level. The limited power available in our optically-powered sensors then leads inevitably to systems with very low center frequencies. For a fixed (optical) pulse width, the power consumption of the optical driver stage is proportional to frequency.

As the carrier frequency is lowered, to conserve power, the filtering problem becomes more difficult to solve. Desired output signals can have frequencies that are an appreciable fraction of the carrier, and of undesired phase detector products.

The demodulator described above solves the filtering problem by greatly reducing the unwanted phase detector outputs. This simplifies the design of the demodulator end of the link, and results in excellent overall system performance.

While first-order designs have been evaluated, and higher orders are possible, in future optical links that use FM, our group is likely to use the Type III zero-order PLL design described above, perhaps combined with a two-pole filter. This seems to have the best combination of performance and simplicity of the FM systems considered.

5 A PERIOD-MODULATING SYSTEM

5.1 Background

A second development is concerned with telemetry of analog information. In telemetry over optical fibers, it happens that representing an analog quantity by means of an analog light level is usually unsatisfactory, and alternative means are used. A commonly used alternative is to employ frequency modulation (FM), since the detection process is then essentially insensitive to the amplitude of the received signal.

Detection (or demodulation) of an FM signal can be accomplished in a number of ways. At high carrier frequencies, a discriminator may be used. Over a wide range of frequencies, a phase lock loop may be used. At low carrier frequencies, better rejection of the carrier is obtained by a system that combines the techniques of state estimation and phase locking to cause a local oscillator to mimic the frequency modulated oscillator. (This last approach was described in Section 2 of this report.)

The present development retains the state estimation to improve the carrier rejection, but it uses a different modulation technique, so that linearization by means of a phase lock loop is not needed.

5.2 Description of telemetry link

There are several ways to implement a voltage controlled oscillator (VCO). At high frequencies, it is usual to control the value of a circuit parameter, such as a capacitance or a mutual inductance, by means of the applied control signal. At low frequencies, it is more common to employ a relaxation oscillator, and to have the control signal control the charging rate of a capacitor in the circuit. In such a system, the capacitor will charge until it reaches a specific target voltage, at which point it is discharged, and the process repeats. If the charging rate is increased, the target is reached sooner, and the frequency is increased. Systems like this can implement very linear frequency modulation.

In the usual phase locked loop application, the modulation is performed by a voltage controlled oscillator (VCO): the output frequency is determined by the input voltage. The period between pulses, the only parameter that can be directly measured using only two pulses, is instead the reciprocal of the frequency. This seemingly intractable mathematical identity has not been resolved in any way that is widely accepted as satisfactory.

The prototype of the new system uses a relaxation oscillator, but the charging rate is held constant and the target voltage is controlled. Now if the target voltage is increased, the capacitor takes longer to reach the value, and the frequency is decreased. By analogy with FM, we may call this

period modulation. The effect is the reciprocal of FM. Figure 4-11 contrasts the two modulation methods.

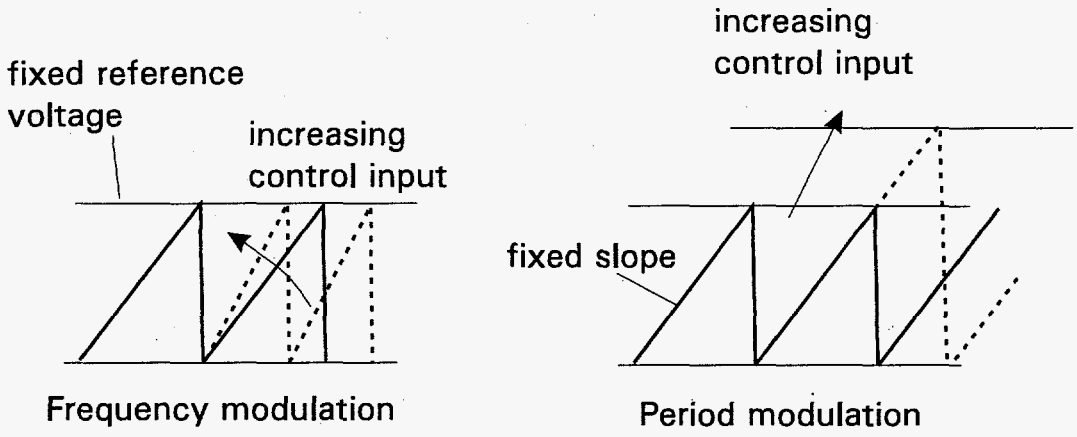


Figure 4-11. Comparison of frequency modulation and "period" modulation

The modulation process can be replicated in the receiver circuit, and the modulating signal recovered efficiently. If the state estimation method of demodulation is used, the output of the sample-hold circuit is an estimate of the modulating signal, without the need to enclose the estimator inside a phase lock loop. Zero- or first-order estimators have been built in our prototypes; higher orders may be possible.

What one would ideally like is a modulating/demodulating system in which it is easy in the demodulator to reproduce what is going on in the modulator. An important result of this new system is that if the period is modulated instead of the frequency, recovery of a high quality estimate of the modulation is greatly simplified.

The prototype modulator is shown in the block diagram Figure 4-12. The system functions as follows. The ramp generator produces a linear ramp with a constant slope. The output of the generator is applied to one input of a comparator. The control voltage, instead of controlling the ramp slope as in FM, is applied to the other input of the comparator. When the ramp voltage reaches the value of the control voltage, the comparator changes state. This causes the ramp generator to be reset, which reverses the state of the comparator. The process repeats, causing a train of period-modulated pulses to be generated.

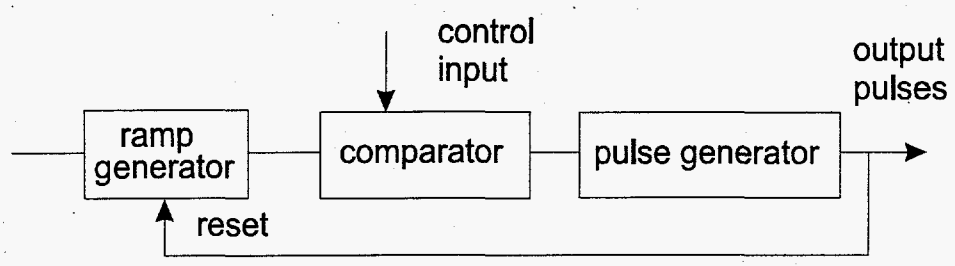


Figure 4-12. Prototype period modulator

The demodulation process is fundamentally a replica of the modulation process. In both the modulator and the demodulator, the pulses correspond to the starting and stopping of the ramp generator. The result is that the ramp generator in the demodulator mimics exactly the ramp generator in the modulator. (This is a clue to the excellent linearity of the system.) As a practical matter, there is a very small delay because of the finite width of the reset pulses.

The information conveyed by the arrival of a pulse from the modulator is precisely that at this instant, the signal voltage was equal to the ramp voltage. Since we know when the previous pulse arrived, and how fast the modulator's ramp generator charges, we know the value of the signal voltage at the instant of each pulse. In fact, since we have a local ramp generator that mimics the one in the modulator, we have a replica of this voltage in the demodulator.

In terms of signal recovery, we can estimate the signal between pulses. The zero-order estimate is that the signal was constant in the interval. A zero order hold circuit operating on a sample of the ramp voltage immediately before the arrival of each pulse implements this. The first-order estimate is that the rate of change of the signal was constant in the interval. This is implemented[3] by a first order hold circuit. In our prototype link we used the sample-hold system described earlier to create the state estimator circuit. The arrangement is shown in block diagram form in Figure 4-13, below.

[3] Approximately. As with the Type III phase locked loop described earlier, the ramp generating circuit that generates the signal that converts a zero order estimate into a first order one is accurate only if the period between pulses is constant. The fact that it is not gives rise to some of the observed nonlinearities of the system.

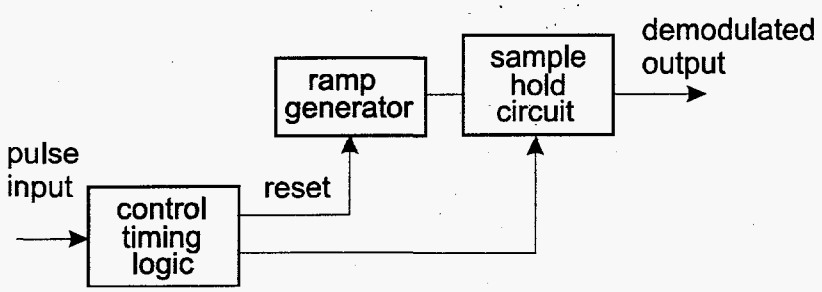


Figure 4-13. Prototype period demodulator

As a practical matter, it is important to ensure that the signal voltage applied to the comparator in the modulator does not reach zero, or more specifically, the discharge value of the ramp. This would force the oscillator to infinite frequency, a practical impossibility. In our implementation, there would be some nonlinearity before this, because of the finite width of the reset pulses, so there is further reason to ensure the condition is not reached. It is a simple matter to offset the control voltage before it is applied to the comparator, and to remove the offset in the demodulator. A bandgap reference diode provides a convenient and temperature stable means of offsetting both modulator and demodulator. The results shown below were obtained with a system of this kind.

An alternative implementation in which the ramp generator operates between the supply rails rather than ground and one rail has not been tested, but it is clear that it would obviate the need to offset the signal. Since the ramp and the signal could each be almost as large as the total supply voltage, it is likely that there would be an improvement in the dynamic range.

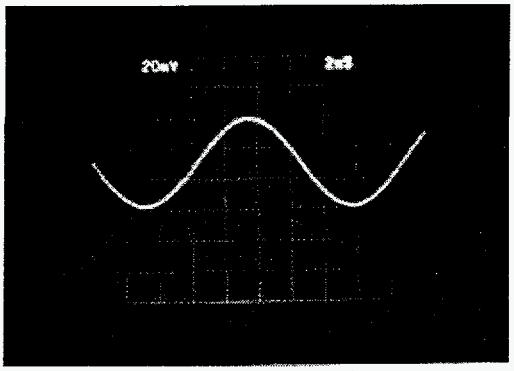
A further small refinement overcomes a possible problem with the modulator. If the reset pulse for the ramp generator is triggered by a change of state in the comparator, a single missed reset can cause the system to "lock up." Once the comparator has gone into the state that indicates the ramp voltage exceeds the input, it will not change state again. If, for some reason, the ramp failed to reset, no further output pulses would be generated. This condition can be avoided by having the reset pulse generated by the state of the comparator, rather than its transition, or by means of a "watchdog" circuit that generates an additional pulse if too much time has elapsed between reset pulses.

5.3 Period modulator - measured results

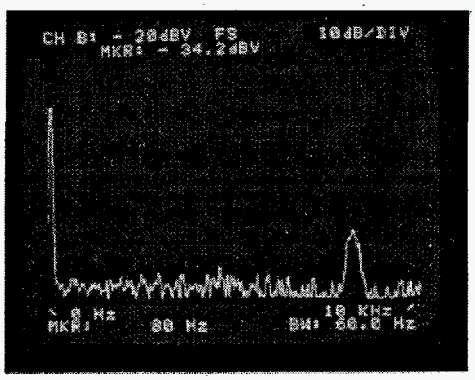
Because the period demodulator uses the state estimator approach, it is to be expected that its performance resembles that of the phase locked loop described in Section 2. This is broadly true, though in some respects it out-performs the PLL system.

Figure 4-14 shows the output signals and the spectra for input signals of 20 mV at 80 Hz. (These results may be compared with Figure 4-8.) In each case the spectrum shows the signal to be 40 dB or more above the fundamental of the carrier. At lower input voltages, the discrete nature of the output is evident. Some uncertainty in the levels of the samples can be seen. This may be due to amplitude noise in the modulator comparator or demodulator sample-hold, or timing jitter.

At higher signal levels there seems to be less evidence of harmonic distortion than in the case of the phase locked loop. Harmonic distortion products appear to be totally absent from the spectrum of the period modulation system.



(a) Output signal



(b) Output signal spectrum

Figure 4-14 (a) and (b). Period modulator/demodulator performance, V_{in} = 20 mV

The dynamic range of the system is shown in Figure 4-15 for a signal frequency of 80 Hz.

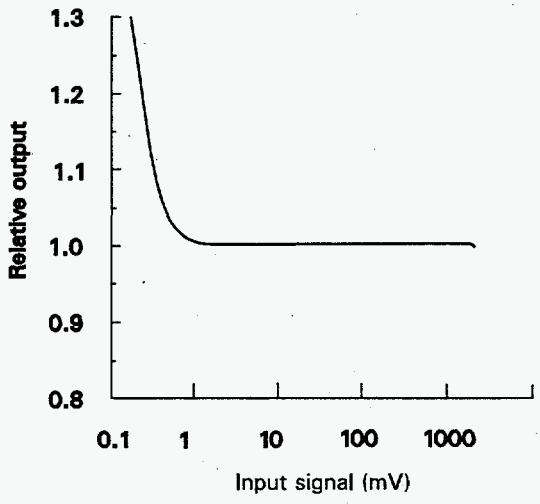


Figure 4-15. Period modulator/demodulator dynamic range

As expected, the dynamic range is comparable with the PLL system, and in the order of 70 dB. The discontinuity in the output near an input of 1 mV is due to the insertion at this signal level of an amplifier (external to the period modulator/demodulator system) to facilitate the measurements.

The frequency response is excellent, as shown in Figure 4-16.

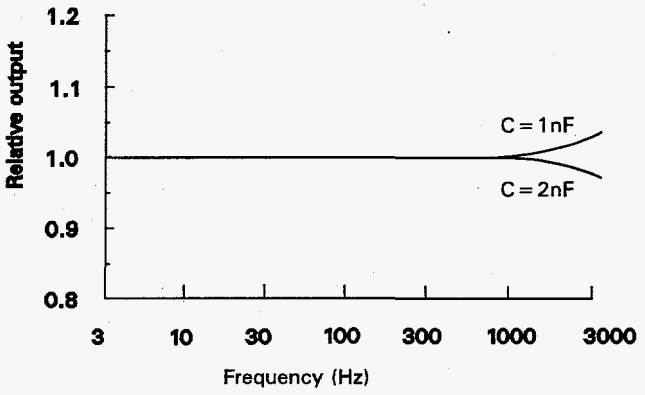
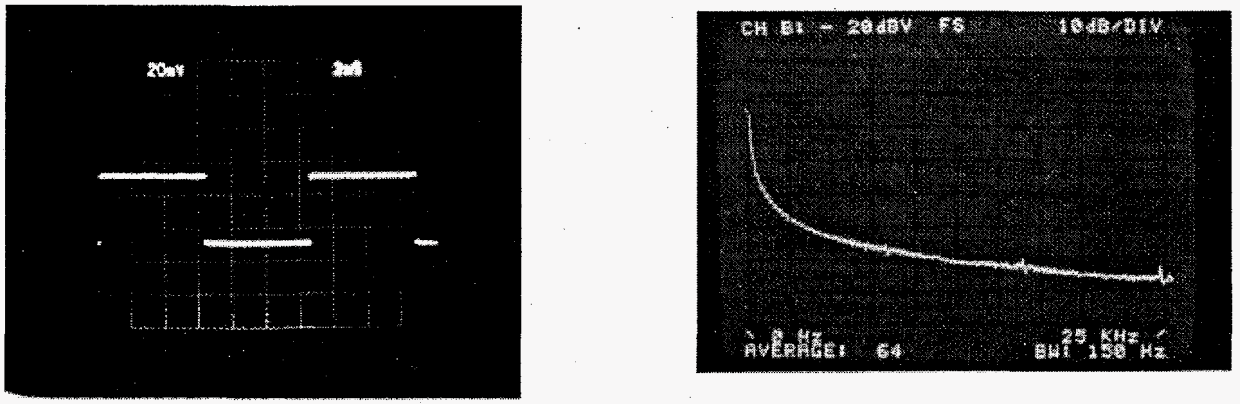


Figure 4-16. Period modulator/demodulator frequency response

Above about 1 kHz there is a slight increase in the output, or a decrease, depending on the size of the capacitor in the receiver sample-hold circuit. This distortion is quite unimportant, as there is in any case a frequency aliasing problem at such high ratios of frequency to sample rate. The frequency response of the period modulated link was not particularly amplitude dependent.

In fact, the link is capable of rather good performance with non-sinusoidal inputs, too. Figure 4-17 shows the system response to a square-wave at 80 Hz.



(a) Output signal

(b) Output spectrum to 25 kHz

Figure 4-17. Square wave response

The spectrum plots are a little misleading, in that the harmonics seem to be asymptotic to some value. Actually, this effect is an artifact of the spectrum analyzer, which has a linear abscissa. The amplitude of the 10-kHz harmonic is 20 dB down on the 1-kHz figure, and the 25-kHz response is 20 dB down on the 2.5-kHz output, indicating the expected 1/f decrease in harmonic output associated with a squarewave.

In some applications of the system as a link, there might be interest in the response to ramp signals. For example, the fields associated with video display terminals (VDTs) may be ramp-like, because of the need to scan the screen. Figures 4-8 and 4-9 show the link response to ramps.

< >modification

The ramp response is clearly as good as the square-wave response. At an input frequency of 1 kHz, there are only 8 or 9 samples per cycle, but within this limitation, the output is a faithful copy of the input. If the system were to be used to measure VDT fields, with a repetition rate of about 15 kHz, a carrier frequency in the order of 100 kHz should suffice for a reasonably accurate measurement of the harmonic energy.

5.4 Concluding Remarks

The period modulator/demodulator system described above enables a telemetry link of low carrier frequency to reproduce low-frequency signals accurately. Harmonic-rich signals, such as might be generated in phase-controlled circuits, and ramp signals, such as those caused by screen scanning circuits, are also transmitted accurately. The system also exhibits good dynamic range without the use of filtering.

A zero-order design has been evaluated, and while higher orders are possible, it may be inferred from the results presented in Section 2 that the extra complexity is not usually warranted by the improvement in performance. In future optical links our group is likely to use the zero-order period modulation system described above in preference even to an FM system using the Type III phase detector.

6 REFERENCES

Adolfsson, M., Einvall, C.-H., Lindbergh, P., Samuelsson, J., Ahlgren, L., and Edlund, H., (1989) "EHV Series Capacitor Banks. A new Approach to Platform to Ground Signalling, Relay Protection and Supervision." IEEE Transctions on Power Delivery, Vol. 4, No. 2, April, pp. 1369-1378.

DeLoach, B.C. and Gordon, E.I., (1981), "Optically powered implant in the human body" Western Electric Technical Digest No. 64, pp. 17-18, October.

DeLoach, B.C., Miller, R.C. and Kaufman, S. (1978), "Sound alerter powered over an optical fiber" Bell System Technical Journal, No. 57, pp. 3309-16, November.

System aspects of tapped dc lines with dynamic controls, J.J. Dougherty and H. Kirkham. IEEE Trans. Power Apparatus and Systems Vol 89 No. 9, Nov/Dec. 1970, pp 2066-2074.

Elgerd, O.I. Control Systems Theory (1967) New York: McGraw-Hill

Gardner, F.M. (1966) Phaselock Techniques, New York: Wiley & Sons.

Hall, P., (1986) "An optically powered sensor network" IEE Colloquium on Distributed Optical Fiber Sensors, (Digest No. 74).

Holte, K.C.(1989), "Technology Requirements for a Competitive Electric Utility Market in the 21st Century" IEEE Power Engineering Review, Vol. 9 No. 7, pp. 18-22, July.

Hopper, A., Temple, S. and Williamson, R. (1986) Local Area Network Design Wokingham, England and Reading, MA: Addison-Wesley.

Kirkham, H. and Johnston, A.R., (1988), Electric and Magnetic Field Meters developed for the U.S. Department of Energy JPL Publication 88-1, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R. and Friend, H.A, (1989), Distribution Automation Applications of Fiber Optics JPL Publication 89-10, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R., Jackson, S.P. and Sheu. K. (1987) AC and DC electric field meters developed for the U.S Department of Energy, JPL Publication 87-20, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R., Lutes, G., Daud, T. and Hyland, S. (1984), Power System Applications of Fiber Optics JPL Publication 84-28, Jet Propulsion Laboratory, Pasadena, California.

Kuo,B.C. Automatic Control Systems Second Edition (1967) Englewood Cliffs, NJ: Prentice-Hall.

McGlade, S.M. and Jones, G.R., (1984) "An optically powered vibrating quartz force sensor" G.E.C. Journal of Research (U.K.) Vol 2, No. 2, pp. 135-8.

Ohte, A., Akiyama, K. and Ohno, I., (1984) "Optically-powered transducer with optical fiber data link" SPIE Vol. 478, Fiber Optic and Laser Sensors II, pp. 33-38.

Schweizer, P., Neveux, L. and Ostrowsky, D.B., (1987) "Optical fiber powered pressure sensor" SPIE Vol. 798 Fiber Optic and Laser Sensors II, pp. 82-5.

Zimmermann, H., (1980) "OSI Reference Model - the ISO Model of Architecture for Open Systems Interconnection" IEEE Transactions on Communications, Volume COM-28, Number 4, April, pp. 425-432.

APPENDIX A.

SOLUTION OF THE TRANSIENT FIELD PROBLEM IN THE TEST CAGE WITH ION CURRENT

A.1 TRANSIENT FIELD

Field measurements in the test cage showed a transient overshoot as the ion source was energized. An example is shown in Figure A.1.

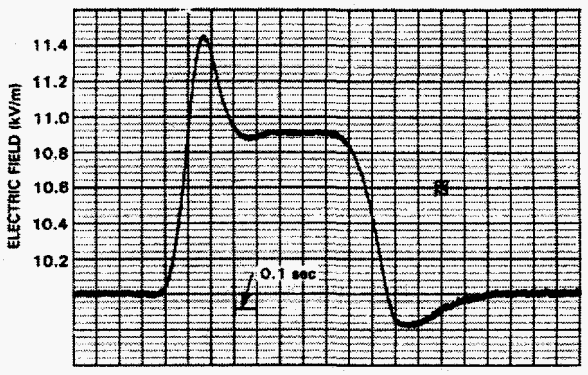


Figure A.1. Example of overshoot of measured field following on source energization

It was reasoned that the overshoot, which persisted for a good fraction of a second, was a genuine result, rather than an artifact of the instrumentation. It seemed reasonable to suppose that the ion current resulted in an increase in the gradient in one part of the cage (and a decrease in another) which eventually disappeared as a steady state was approached. In particular, it was thought that when only part of the cage contained ions, the ionized part of the cage would have less voltage across it, and the non-ionized part more, than when the entire cage was devoid of ions, or when the ions completely filled the cage. It will be shown that these initial ideas were not quite correct.

To analyze this problem, the cage can be divided into two regions, above and below a moving boundary. Above the boundary, ions are present, and the distribution of voltage and gradient can be described by Poisson's equation. Below the boundary there are no ions, and the conditions for Laplace's equation exist.

Figure A.2 shows the cage with the ionized region part way down. The effect of diffusion is neglected. It is assumed that the front edge of the ionized region remains sharply defined as it moves into the cage.

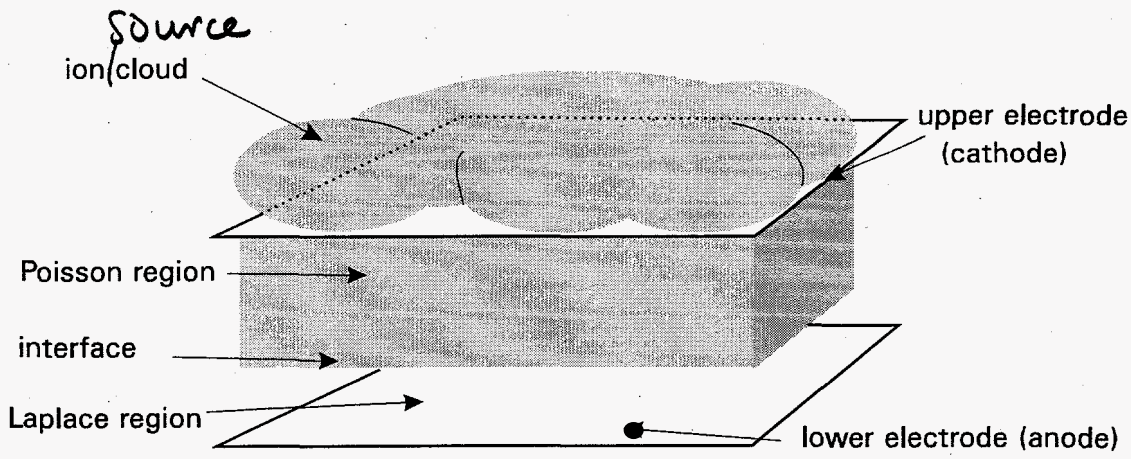


Figure A.2. Test cage with ionized region

In the Poisson region above the boundary, Poisson's law in rectangular coordinates

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \frac{-\rho}{\epsilon} \tag{A.1}$$

can usually be reduced to the complete differential equation

$$\nabla^2 V = \frac{\partial^2 V}{\partial z^2} = \frac{-\rho}{\epsilon} \tag{A.2}$$

for the one dimensional case, in view of the x and y symmetry. Here V is the potential, ρ the space charge density and ϵ the permittivity of free space. The cloud of charge is assumed to be moving in the -z direction. Until the steady state is achieved, the equation should in fact show the variables also as a function of time:

$$\frac{\partial^2 V(z,t)}{\partial z^2} = \frac{-\rho(z,t)}{\epsilon} \tag{A.3}$$

In addition to Poisson's equation, A.3, there are constitutive relationships describing the behavior of the charge. The current in the cage is given by the sum of two components, the conduction current, which exists only in the Poisson region and the displacement current, which is necessary to avoid violation of Kirchoff's law at the boundary. Thus

$$J_{total} = J_{cond} + J_{displ} \tag{A.4}$$

where J_{cond}, the conduction current density, is given by the product of charge density and carrier velocity

$$J_{cond}(z,t) = \rho(z,t)\, v(z,t) \tag{A.5}$$

and the displacement current, J_{displ}, is given by the time rate of change of the electric field

$$\begin{aligned} J_{displ}(z,t) &= \epsilon \frac{\partial E(z,t)}{\partial t} \\ &= \epsilon \frac{\partial^2 V(z,t)}{\partial z\, \partial t} \end{aligned} \tag{A.6}$$

The total current is therefore given by

$$J(z,t) = \rho(z,t)\, v(z,t) + \epsilon \frac{\partial^2 V(z,t)}{\partial z\, \partial t} \tag{A.7}$$

Note that the charge carriers move with a velocity v given by the magnitude of the field E:

$$v(z,t) = \mu E(z,t) = -\mu \frac{\partial V(z,t)}{\partial z} \tag{A.9}$$

where μ is the parameter termed mobility, assumed constant. (This expression is equivalent to the statement that energy is not conserved, and it may be derived from energy considerations of a nonconservative system.)

Therefore,

$$J(z,t) = \epsilon \frac{\partial^2 V(z,t)}{\partial z\, \partial t} - \rho(z,t)\, \mu \frac{\partial V(z,t)}{\partial z} \tag{A.9}$$

This can be rearranged to yield an expression for ρ(z,t):

$$\rho(z,t) = \frac{\left(\epsilon \frac{\partial^2 V(z,t)}{\partial z\, \partial t} - J(z,t) \right)}{\mu \frac{\partial V(z,t)}{\partial z}} \qquad \text{(A.10)}$$

Substitution of this value for $\rho(z,t)$ into equation A.3 (Poisson's equation) yields

$$\epsilon\mu \left(\epsilon \frac{\partial V(z,t)}{\partial z} \right) \left(\frac{\partial^2 V(z,t)}{\partial z^2} \right) + \epsilon \frac{\partial^2 V(z,t)}{\partial z\, \partial t} = J(z,t) \qquad \text{(A.11)}$$

A number of workers have investigated the question of gaseous conduction[1]. It transpires that equation A.11 has other applications : we are indebted to Professor W. J. Gajda of the University of Missouri at Rolla for pointing out the similarity of our problem to the problem of current injection in solids.

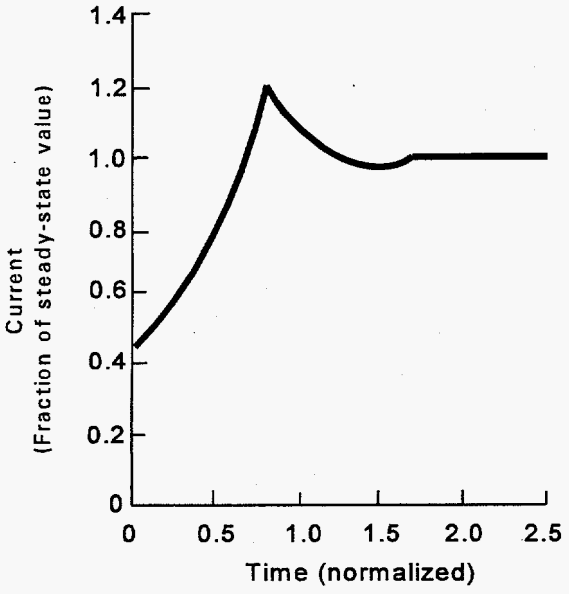
The solution to the problem of SCL injection in solids is given by Lampert and Mark (1970)[2]. There it is shown that the injection current increases to a peak value about 21% above the final steady state value by the time the front edge of the charge reaches the anode. From this value, there is a decrease in current, with a slight undershoot. The time taken for the leading edge of the charge to reach the anode is about 78% of the time calculated for the transit of a single ion in an otherwise Laplace region.

The current waveform after Lampert and Mark is shown in Figure A.3, where it is compared with the field measured by the probe in the JPL test cage. The current curve is normalized; time and field scales are given for the cage field results. In the examples shown, the anode-cathode spacing was 60 cm, and the applied voltage was -6 kV.

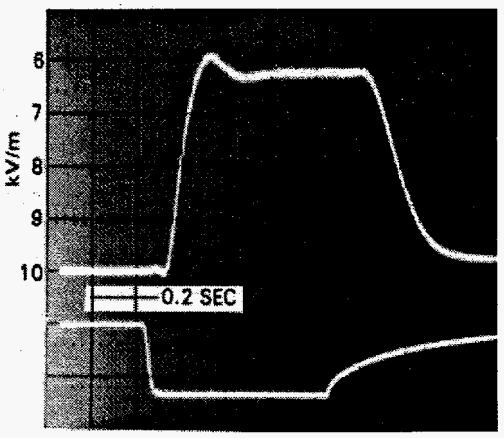
---

[1]Zahn, Tsang and Pao (1974) present an analysis of unipolar ion conduction which includes a transient similar to the one we observed. Their paper presents its results chiefly in the form of computer-generated graphs, although much of the analysis is closed form. The work is extended by Zahn and Pao (1975) and by Zahn, Pao and Tsang (1976). A model for bipolar conduction is developed by Zahn (1975). The problem is analyzed for spherical and cylindrical geometries with unipolar conduction by Zahn (1976). Crow, Auer and Allen (1975) discuss the expansion of a plasma into a vacuum, but the plasma is collisionless.

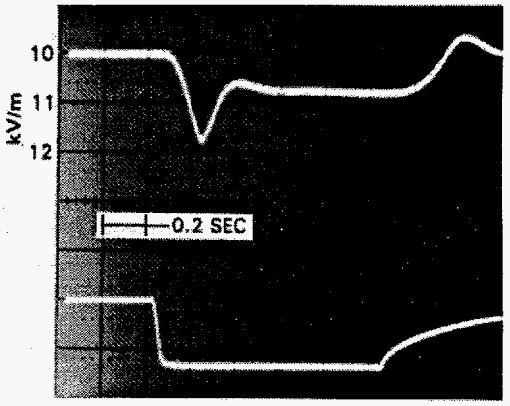
[2]The question of the transient in solids is also examined by Many and Rakavy (1962) for SCL currents both with and without trapping.



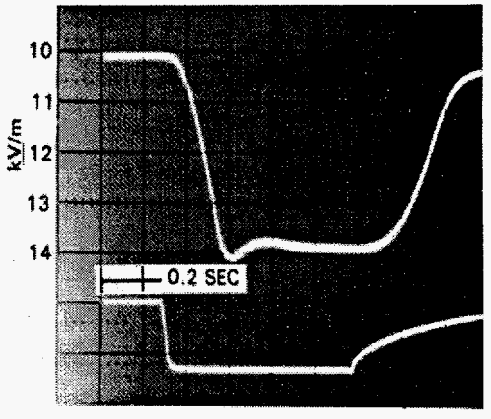
(a) Current transient
(after Lampert and Mark)



(b) Field near cathode



(c) Field at center of cage



(d) Field near anode

Figure A.3. Field transient in test cage

The cage results show that, as expected, the field near the cathode decreases when the ion source is energized. The field at the anode increases. The cage center is near the point at which the field changes little (in the final steady state) as a result of the ion current. Nevertheless, the transient

is still clearly visible.

The lower trace in the oscilloscope pictures (Figure A.3; b, c and d) is a marker to show when the ion current source was energized. In each case, there seems to be a small delay between the energizing of the ion source, and the appearance of the transient in the field. This may be due to the transit time of the ions from their true source to the cathode. The delay is no different whether the probe is near the anode or the cathode. The calculated transit time for an isolated ion in the cage under these conditions is 0.4 s, (assuming an ion mobility of 1. 5×10^{-4} m^2/V s) so that the current peak should be reached after about 0.31 s. This is in good agreement with the measured results.

It is clear that most (if not all) of the transient increase in the field can be attributed to the increase in the current. The rounding of the field waveform is due in part to diffusion effects, and in part to the filtering of the output signal from the field meter. Apart from this, the comparison is clearly very good.

A. 2 Space Charge Limited Current

The question arises, is there a way to calculate the value of the steady-state current in the cage, to "denormalize" the curve given by Lampert and Mark? If so, the calibration of the field meter would no longer rely on the measurement of current in order to know the field in the cage when there was ion current.

The steady-state problem is soluble in closed form. By taking only the steady-state part of Equation A.11, we obtain

$$\epsilon \mu \left(\epsilon \frac{dV(z)}{dz} \right) \left(\frac{d^2V(z)}{dz^2} \right) = J(z) \qquad \text{(A.12)}$$

Simplifying,

$$\epsilon \mu E \frac{dE}{dz} = -J \qquad \text{(A.13)}$$

so that

$$\frac{dE}{dz} = \frac{-J}{\mu \epsilon E} \qquad \text{(A.14)}$$

integration gives that

$$\int E\,dE = \int \frac{-J}{\mu \epsilon} dz \qquad \text{(A.15)}$$

or

$$E^2 = \frac{-2Jz}{\mu\epsilon} + C \tag{A.16}$$

where C is a constant of integration.
Therefore

$$E = \pm\left(C - \frac{2Jz}{\mu\epsilon} \right)^{1/2} \tag{A.17}$$

(It will be seen later that the negative root is required for a solution.)

This can be integrated by means of a substitution. First, we simplify by letting

$$a = \frac{2J}{\mu\epsilon} \tag{A.18}$$

so that Equation A.17 becomes

$$E = \pm(C - az)^{1/2} \tag{A.19}$$

Next, we let

$$F^2 = C - az \tag{A.20}$$

so that

$$2F\,dF = -a\,dz \tag{A.21}$$

and

$$V = -\int E\,dz = -\int \pm(C - az)^{1/2}\,dz \tag{A.22}$$

At this point it is clear that the negative root must be used if V is to be positive. Thus

$$
\begin{aligned}
V &= \int F \frac{2F\,dF}{a} \\
&= \frac{2F^3}{3a} + K \\
&= \frac{2}{3a}(C - az)^{3/2} + K \\
&= \frac{\mu\epsilon}{3J}\left(C - \frac{2J}{\mu\epsilon}z\right)^{3/2} + K
\end{aligned}
\tag{A.23}
$$

where K is another constant of integration.

K can be evaluated directly. Since we have $V=0$ when z=0, it follows from A.23 that

$$K = \frac{\mu\epsilon}{3J}C^{3/2} \tag{A.24}$$

whence

$$V = \frac{\mu\epsilon}{3J}\left(\left(C - \frac{2J}{\mu\epsilon}z\right)^{3/2} - C^{3/2}\right) \tag{A.25}$$

It is an experimentally observed fact that the field decreases near the source of the ions, and is zero at the boundary in the space-charge-limited case. (This zero field is equivalent to the definition of space charge limiting.) This being the case, $E=0$ at z=h, so that the value of the constant of integration C can be found directly from Equation A.17:

$$C = -\frac{2J_{SCL}h}{\mu\epsilon} \tag{A.26}$$

(J is negative in this coordinate system.)

We can also use equation A.25 for voltage to apply a boundary condition that at z=h, $V=V_h$. In the SCL case the term inside the parentheses (which represents the gradient) vanishes, so that

$$V_h = -\frac{\mu\epsilon}{3J_{SCL}}C^{3/2} \tag{A.27}$$

whence

$$C = -\left(\frac{3 J_{SCL} V_h}{\mu \epsilon} \right)^{2/3} \tag{A.28}$$

The two values of C, in A.26 and A.28, must be the same, ie

$$\frac{2 J_{SCL} h}{\mu \epsilon} = \left(\frac{3 J_{SCL} V_h}{\mu \epsilon} \right)^{2/3} \tag{A.29}$$

$$\left(\frac{2 J_{SCL} h}{\mu \epsilon} \right)^3 = \left(\frac{3 J_{SCL} V_h}{\mu \epsilon} \right)^2 \tag{A.30}$$

$$J_{SCL} = \frac{3^2 V_h^2 \mu \epsilon}{2^3 h^3} \tag{A.31}$$

We therefore arrive at the interesting result that the value of current at which space charge limiting occurs can be calculated in a fairly straightforward way once the size of the cage and the applied voltage are known.

For example, during the tests at NBS we had

$V = 1006$ V

$\mu = 1.5 \times 10^{-4}$ m^2/V sec

$\epsilon = 8.854 \times 10^{-12}$ F/m

$h = 0.17$ m

from which a value for the SCL current can be found. The result is approximately 307 nA. This compares well with the current actually observed, 303 nA.

Equation A.31 can serve as a simple way to measure the mobility of the charge carriers. Indeed, it is likely that the NBS value for the ion mobility used in the tests in Gaithersburg was obtained by recasting equation A.31 to yield a value for μ.

A. 3 Gradient in a Poisson field

So far, we have considered the voltage and current conditions in the cage. The equations

developed can also be used to find the gradient as a function of height, assuming space-charge-limited conditions. Thus, from Equation A.16 and assuming $J = J_{SCL}$:

$$E = \pm\left(C - \frac{2Jz}{\mu\epsilon} \right)^{1/2} \tag{A.32}$$

From A.26 we have

$$C = -\frac{2Jh}{\mu\epsilon} \tag{A.33}$$

so that

$$\begin{aligned} E &= \pm\left(\frac{2Jh}{\mu\epsilon} - \frac{2Jz}{\mu\epsilon} \right)^{1/2} \\ &= \pm\left(\frac{2J}{\mu\epsilon} \right)^{1/2} (h - z)^{1/2} \end{aligned} \tag{A.34}$$

It is possible to eliminate J, since Equation A.31 gives that

$$J_{SCL} = \frac{3^2 V_h^2 \mu\epsilon}{2^3 h^3} \tag{A.35}$$

whence

$$\begin{aligned} E(z) &= \pm\left(\frac{2 \times 3 V_h^2}{2^3 h^3} \right)^{1/2} (h - z)^{1/2} \\ &= -\frac{3}{2} \frac{V_h}{h} \left(\frac{h - z}{h} \right)^{1/2} \end{aligned} \tag{A.36}$$

so that the Poisson field can be expressed in terms of the Laplace value V_h/h in the form

$$E_{Poisson} = \frac{3}{2} H(z) E_{Laplace} \tag{A.37}$$

where H is a normalized height factor $\sqrt{(h - z)/h}$

Equation A.37 shows that the maximum field intensification due to ion current in a planar cage is 50%, and as expected the field drops to zero at the cathode.