

THE BIRTH AND EVOLUTION OF SUNSPOTS: OBSERVATIONS

Patrick S. McIntosh
Space Environment Laboratory
National Oceanic and Atmospheric Administration
Boulder, Colorado 80303

The structure, dimensions and dynamics of sunspots and sunspot groups are reviewed as they are observed in white light. The fine structure of sunspots during their formation and dissolution suggests an intimate relationship with the photospheric granulation. The distribution of sizes among the smaller spots in sunspot groups shows a maximum at a diameter of 5000 km, corresponding to the newly-announced mesogranulation. The position of the initial spots of a group, and the scale and distribution of spots within a mature group, reflect the next largest scale of solar convection, supergranulation. The growth of the larger sunspot groups proceeds by the successive emergence of up to six bipolar sunspot "sets" that appear at approximately 24-hour intervals. The spot-group axis typically begins with a high inclination to the solar equator and rotates to a small inclination through the addition of new spots. The direction of group-axis rotation appears to obey a Coriolis law for outflowing gas. The rotation of clusters of spots in spot groups obeys a Coriolis law for inflowing gas. Large spots form through the coalescence of small spots.

The formation of the larger sunspot groups is not random. Large and active sunspot groups are observed to form repeatedly on the same boundary of a large-scale magnetic-field pattern over a period of many months. Large sunspot groups often form sequentially to the east, following the slow rate of motion of large-scale features adjacent to the sunspot group and lying at higher latitude. Long-lived spots form on the low-latitude boundary of large-scale magnetic patterns, and share their slow rate of solar rotation. Great active regions form in areas of maximum shear among the large-scale patterns. Sunspots are small and short-lived in longitudes of large-scale divergence.

1. Introduction

This review draws upon the author's own 25 years of sunspot observations and studies of their relationships to other forms of solar activity. Much of this point of view was developed while participating in real-time solar-terrestrial forecasting, which requires daily

examination of every active region on the sun. The close correlation between sunspot behavior and flare activity has made sunspot research a natural first step in efforts to develop effective methods for predicting solar flares. This research has drawn upon films being accumulated daily by the 15 cm white-light patrol telescope at the Sacramento Peak National Observatory. This is perhaps the only patrol in existence which photographs the entire solar disk as often as most flare patrols, providing a complete record of the birth and evolution of all sunspot groups since September, 1963 (McIntosh, 1964). Photospheric details to one-arc-second resolution are enlarged to a useful scale through the making of internegatives at 16 times the scale of the original patrol film (McIntosh, 1972a). Only about 300 of the nearly 7500 sunspot groups recorded by this patrol have been studied so far; but, these examples already demonstrate how incomplete the harvest of white-light sunspot observations has been over the past 370 years.

The complex and rapidly-evolving sunspots associated with high flare activity are hardly suitable for building classic physical models of sunspots. The theoretician's sunspot is the isolated, large, stable and symmetric spot; however, such spots are the exception, and they must be viewed as the remaining "leader" spot of a bipolar spot group that once had many other spots. The birth and evolution of leader sunspots differ from other spots in the group; therefore, a physical model based only on properties of leader spots may miss critical aspects of the general sunspot problem.

In reviewing the structures, dimensions, dynamics and distributions of sunspots over the solar surface associations with the four basic scales of solar convection are inferred: granulation, mesogranulation, supergranulation and giant cell convection. The hydrodynamics of vorticity, shear, convergence and global circulation appear to be fundamental ingredients in the physics of sunspots. The view of sunspots on the larger scale of global magnetic fields and over periods of years has developed from a program of mapping solar activity together with large-scale magnetic-polarity patterns (H-alpha synoptic charts) for each solar rotation since 1964 (McIntosh, 1972b, 1972c, 1979).

2. White-light Sunspot Structures

Excellent white-light photographs of the sun (Figure 1) resolve a wealth of fine structure that indicates variations in temperature and magnetic fields within the photosphere. Sunspots are composed of a dark umbra, distinctly darker than the spaces among the photospheric granules, and some of the larger umbrae are partially, or entirely, surrounded by a lighter penumbra. The umbrae are sites of primarily vertical magnetic fields and the penumbrae occur where there are strong horizontal magnetic fields.

2.1 Pores

The distinction between pores and sunspots was made confusing by Bray and Loughhead (1964) in that they defined pores as any spot lacking penumbra. More than half the spot groups never develop penumbra, and

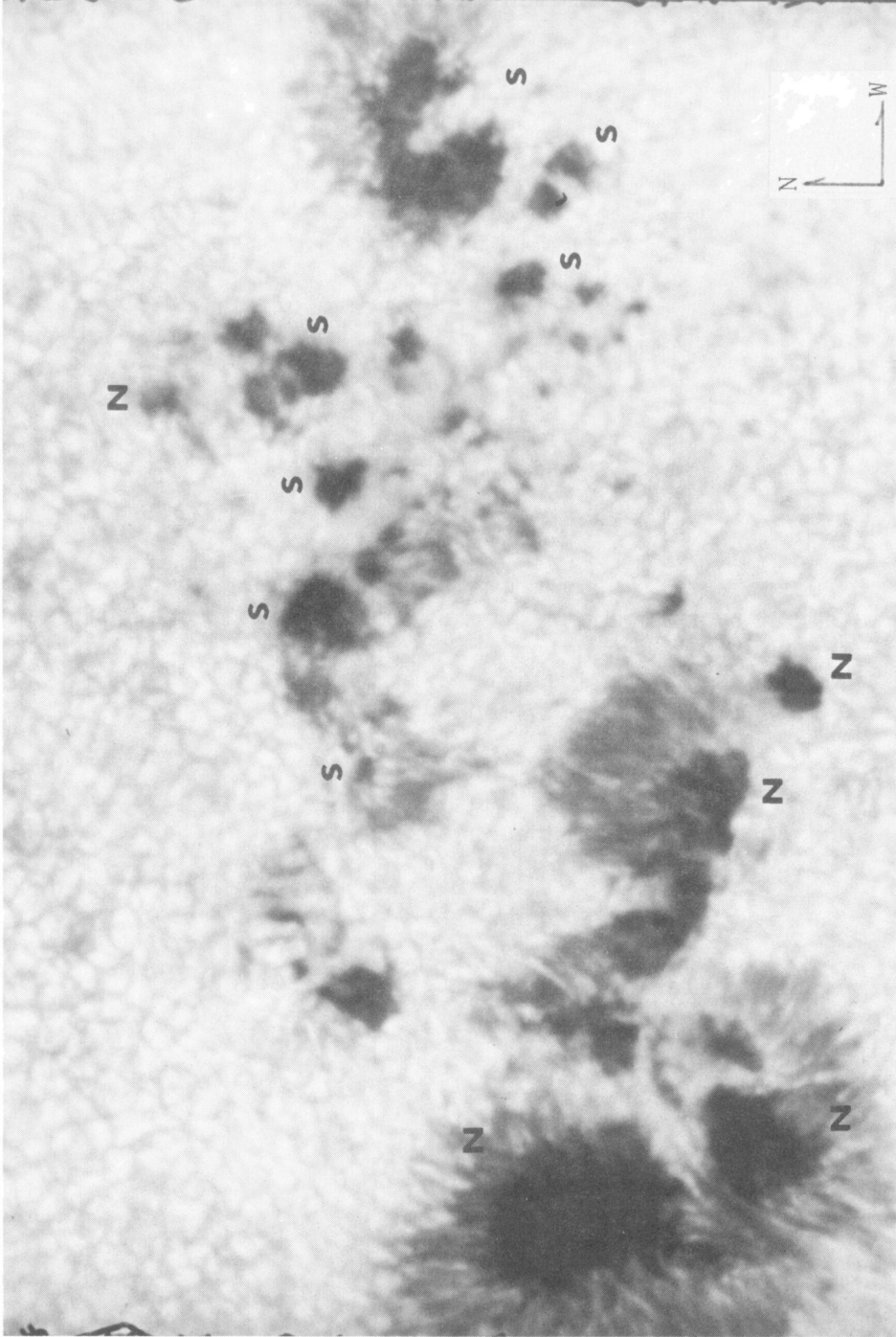


Figure 1. Photospheric granulation, pores, spots, penumbra (rudimentary and mature) and a light bridge all appear together in the finest white-light photograph obtained from Sacramento Peak Observatory up to 14 April 1963, 14:02:40 U.T. Photographed on Kodak High Contrast Copy film with the 12-inch (30 cm) horizontal telescope. The scale is 0.6"/mm; resolution is better than 0.5" arc.

most of the spots within any given group are not involved with penumbra. The correct definition of a pore, continuing the tradition of observers over the past century or more, is: darkening in the photosphere similar in intensity to the intergranular spaces and with a diameter up to 2500 km. While agreeing with this definition, Simon and Weiss(1970) add that the pore has a magnetic-field strength in excess of 1500 gauss.

The lifetime of a pore may be a function of the kind of region in which it occurs. In the quiet photosphere pores have lifetimes of 10-15 minutes, similar to the granulation. Many of them near sunspot groups are also short-lived, but others are observed to persist up to an hour. Some regions may form pores repeatedly in the same area for more than a day. A pore can be the initial stage in forming a spot, with the transition being one of a drop in intensity rather than a coalescing of two or more pores (Schröter, 1962). Visual observers often add pores to their spot counts if seeing conditions are especially good. Two examples of pores appear near the north edge of Figure 1. The one at the left (east) had disappeared two minutes after this frame was obtained.

2.2 Umbrae = Spots

The count of sunspots that goes to make up the index known as the sunspot number is a count of separate umbrae, and includes counting each of the separate umbrae within a single area of penumbra. The definition of spot as an umbra with penumbra would lead to a drastic reduction in the sunspot count, if this definition were adopted.

An umbra is distinguished from a pore by intensity. The high resolution of Figure 1 permits the detection of spots that are smaller than most of the granules in the photosphere adjacent to the spot group; the smallest appears to have a diameter near the 300 km limit of resolution. It seems to be no more than an especially dark intergranular space.

All umbrae have irregular outlines, with the smaller spots having the coarser irregularities on a scale seemingly determined by the size of the adjacent photospheric granules. The boundary between umbra and penumbra in large spots is serrated. The dark prongs correspond to the inner base of dark penumbral filaments and the notches correspond to the base of bright penumbral filaments. The distance between dark penumbral filaments (and between umbral "prongs") is about 500 km.

Umbrae are non-uniform in intensity, with changes in intensity occurring in cellular units. The cell size in the smaller spots is similar to the granules in the surrounding photosphere. At this point it should be pointed out that the size of the granulation becomes reduced in areas interior to strong sunspot groups such as the one in Figure 1. The granulation is distinctly smaller in the broad corridor of photosphere between the large mass of spots and the numerous small spots in the leading (western) portion of the group. This corridor coincides with the line of magnetic polarity inversion as inferred from simultaneous H-alpha spectroheliograms.

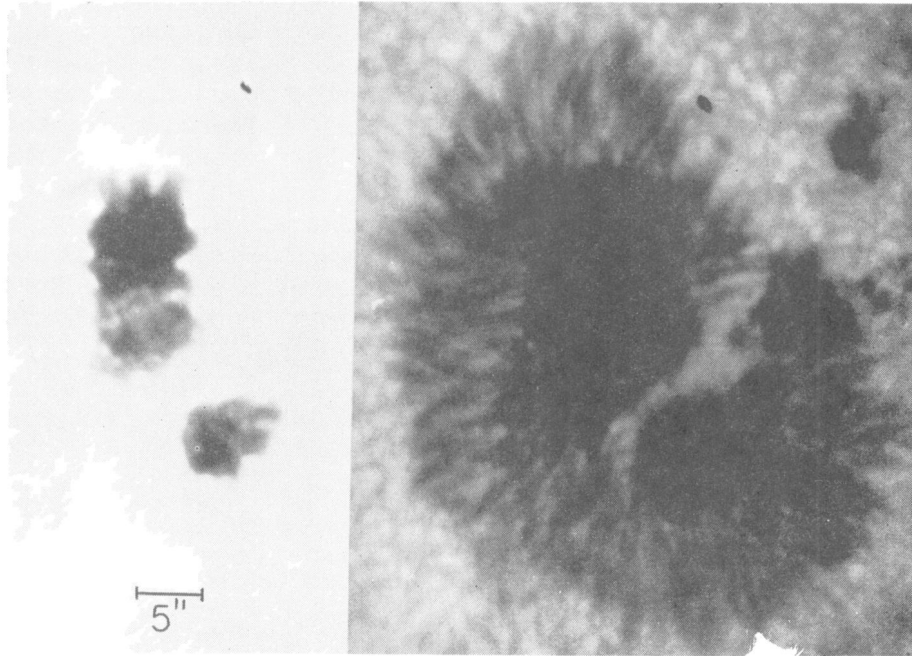


Figure 2. Two prints from the same negative of 13 April 1963, 14:13:20 U.T. taken with the 12-inch (30 cm) horizontal telescope. At left, the negative was printed to show chains of tiny emission points within the sunspot umbra. These chains are aligned with bright penumbral structures that extend from the chains to the right, as shown in the print at right exposed to show penumbral and photospheric structure. Reproduced from McIntosh (1964).

The larger spots, which usually have much stronger magnetic fields, exhibit interior structuring on a smaller scale, as if the magnetic fields reduce the scale of convection. Seldom are the intensity contours within an umbra symmetric with the outer boundary of the umbra (e.g. Figure 2). Most large umbrae are composed of two or more intensity minima, and these appear to correspond to the separate small umbrae that coalesced to form the large spot (see Section 3.1).

In addition to the cellular intensity variations within umbra there are umbral dots. These bright features are well shown in Figure 2, a photograph of the large spot in Figure 1 one day earlier. This photograph was exposed longer than normal for photospheric structure, yet the film's dynamic range permitted printing to show either the dots or the surrounding photospheric granulation. These dots are mere points on even this fine negative, indicating an upper limit to their diameter at 300 km. The dots formed two chains across the umbra, meeting the umbral boundary at the base of bright penumbral filaments.

Much attention has been paid to effects of scattered light on measurements of the intensities of sunspot umbrae (Zwaan, 1965, 1974; Mattig, 1971), and correctly so. However, it would appear that the conclusion that there is no more than a 15% variation in intensity among all sunspots is a misleading statement. Almost all the detailed photometric studies have been conducted on large leader sunspots, so the conclusion is valid only for this set of spots. The Sacramento Peak full-disk patrol films include many long-exposure negatives obtained

during good seeing. These show large variations in intensity among the spots within individual spot groups, and the relative intensities remain unchanged from frame to frame, eliminating the possibility that seeing effects caused the differences in intensities. The spot intensities are, on the average, lower for the larger umbrae; but, there are frequent exceptions. Sunspot intensity is not well-correlated with either spot area or the presence of penumbra. It is clear from the original negative that the leader spot in Figure 1 is much brighter than the smaller spot in the north-central portion of the group. Note, too, that the presence of penumbra and its extent from a spot may not correlate with the size and intensity of the umbra.

Some important umbral dimensions:

Smallest spot: < 300 km diameter

Spots with lifetimes > 2 days: 5000 km diameter (see Section 2.5)

Maximum diameter without penumbra: 11,000km (16 August 1963)

Minimum diameter with penumbra: 2500 km

Maximum diameter single umbral area: 23,000 km (20 March 1966)

2.3 Penumbra

The grey penumbral areas surrounding the larger sunspots occur in four categories of structure: rudimentary (formative stage); rudimentary (dissolving stage); mature, radial penumbra; and thick, dark penumbral filaments in regions of strong transverse magnetic field and high longitudinal gradient. The first three are present in Figure 1.

The indistinct penumbra interconnecting small spots in the north-central portion of the group in Figure 1 is rudimentary in a formative stage. Portions of this penumbra appear as elongated photospheric granules with especially dark intergranular spaces. This type of penumbra was first described by Bray and Loughhead (1964). Included in the same rudimentary category, but with slightly different structure, is the penumbra on the leader spot (extreme west end of group). This penumbra had disappeared by the next day. The structure is more granular than filamentary and the outer edge is strongly scalloped by intruding photospheric granules. Both types of rudimentary penumbra are brighter than mature filamentary penumbra and seldom extend completely around a spot. The distance from umbra to edge of rudimentary penumbra is 4000 km in the case of this leader spot, but is usually half this value. Rudimentary penumbra is difficult to discern with visual observations at the small scales typical of daily sunspot drawings.

Mature penumbra is what is most commonly observed surrounding large leader spots. Its filamentary bright and dark structures radiating from the umbra were best documented in the high-resolution Stratoscope photographs (Danielson, 1961). The penumbra surrounding the largest spot and adjacent umbrae in Figure 1 is mature. It extends 6000 km from the umbral boundary at its widest. Larger spots than these can have penumbra extending 12,000 km from the umbral boundary and a total spot diameter as much as 52,000 km.

Penumbra on the interior and follower spots of a spot group is often observed extending from only part of the umbral boundary. In the example in Figure 1, the penumbra in the south-central portion of the group extends toward penumbra of opposite magnetic polarity across the corridor of photosphere. This indicates relatively strong transverse magnetic fields across the line of polarity inversion. The evidence like this leads to the conclusion that the presence of, and distribution of, penumbra is not a function of the size and/or intensity of a spot but depends more strongly on the overall geometry of the sunspot-group magnetic fields. In many of the examples in this review there are groups with several spots of equal size but with only a few surrounded by penumbra.

The fourth type of penumbra occurs only in those rare spot groups with large and rapidly-evolving spots of opposite polarity involved in a common penumbra - the magnetic "delta" configuration (Figure 3). The dark penumbral filaments in the interior of the group are wider and darker than normal penumbra. The filaments are oriented parallel to the borders of umbrae of opposite polarity on either side of the area of dark penumbra; i.e. the penumbral structure parallels the line of polarity inversion. A reasonable interpretation of this structure is that it represents exceptionally strong transverse magnetic fields that have been highly sheared from a potential form. This is a favorable situation for flare production. Similar occurrences to Figure 3 occurred on July 7, 1966 (McIntosh, 1969a), May 23, 1967 (McIntosh, 1969b) and August 4-7, 1972 (McIntosh, 1973). All four examples had large solar flares occur over these areas of dark penumbra, and the flares were accompanied by strong proton emission.

It is common to have a scattered arrangement of small spots within an extensive area of complex penumbra, with the ratio of penumbral area to umbral area much higher than in classic leader sunspots. This ratio can actually extend toward infinity by the disappearance of all the umbrae, leaving an isolated area of pure penumbra! Such occurrences are not unusual in very complex sunspot groups. The physical interpretation of such patches of penumbra is that relatively strong transverse magnetic fields occurred near the photosphere with little or no areas of vertical fields. Penumbrae without interior umbrae seldom persist more than one day.

The process of forming and dissolving penumbra is well illustrated by Bumba (1965a). The formation of penumbra is preceded by the appearance of chains of granules aligned radial to the sunspot and a darkening of the intergranular lanes alongside these chains. The dissolution process occurs by the encroachment of more and more photospheric granulation into the borders of the penumbra, while the penumbra itself evolves from filamentary to granular fine structure. The intensity of the penumbra increases during this evolution.

2.4 Sunspot Group Classification

No two sunspot groups are exactly alike, and the differences tend to become greater in larger spot groups. On the other hand, there is a finite range of sizes and structures among sunspot groups so that most

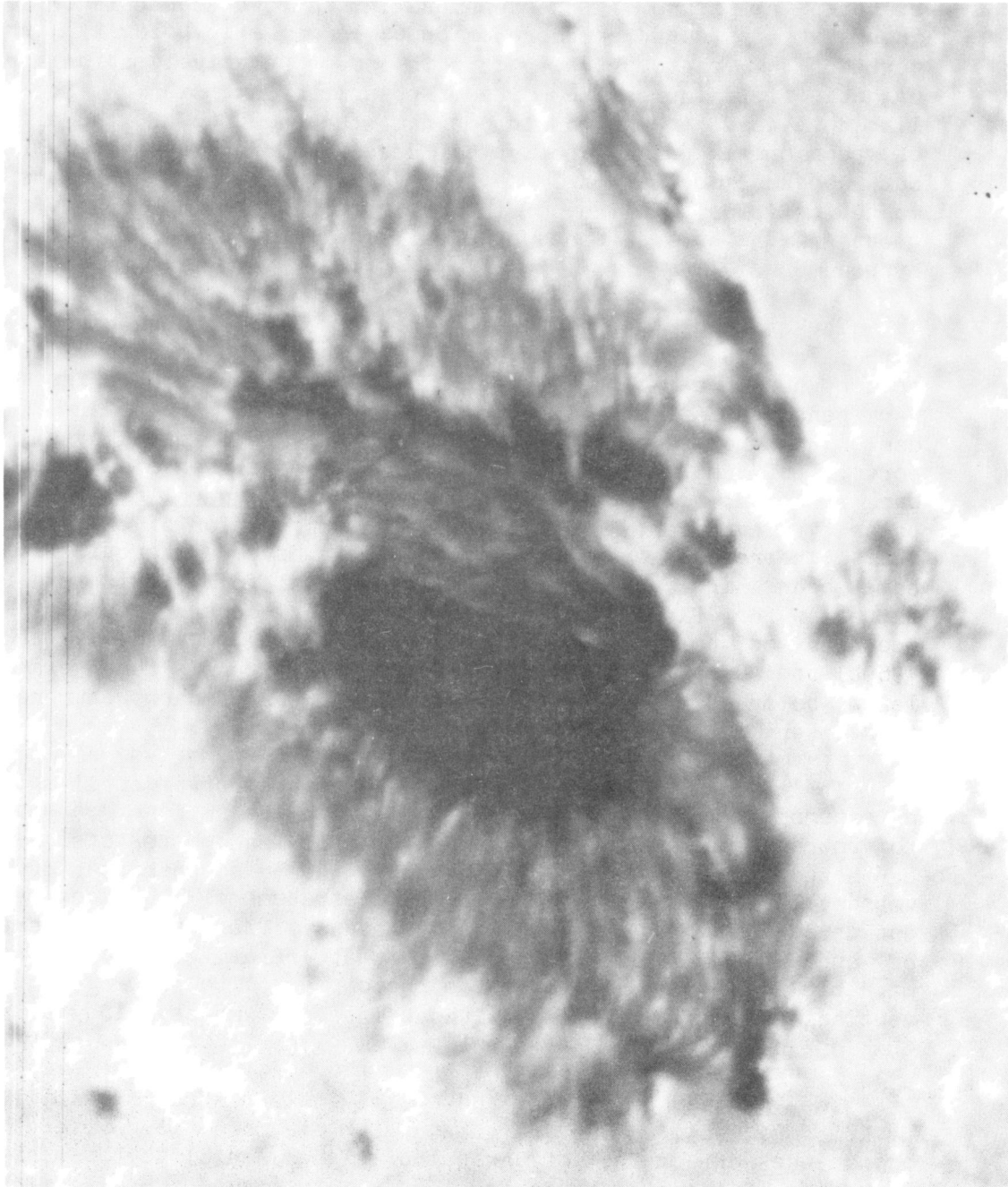


Figure 3. Exceptionally dark and thick penumbral structure occurred in the area between strong spots of opposite magnetic polarity on 24 September 1963, 14:40:15 U.T. The inclusion of strong spots of opposite polarity within a common penumbra is a "delta" configuration. Photographed with the Sacramento Peak 12-inch (30 cm) horizontal telescope. The scale is 0.4"/mm.

Modified Zurich Class

McINTOSH
SUNSPOT GROUP CLASSIFICATION

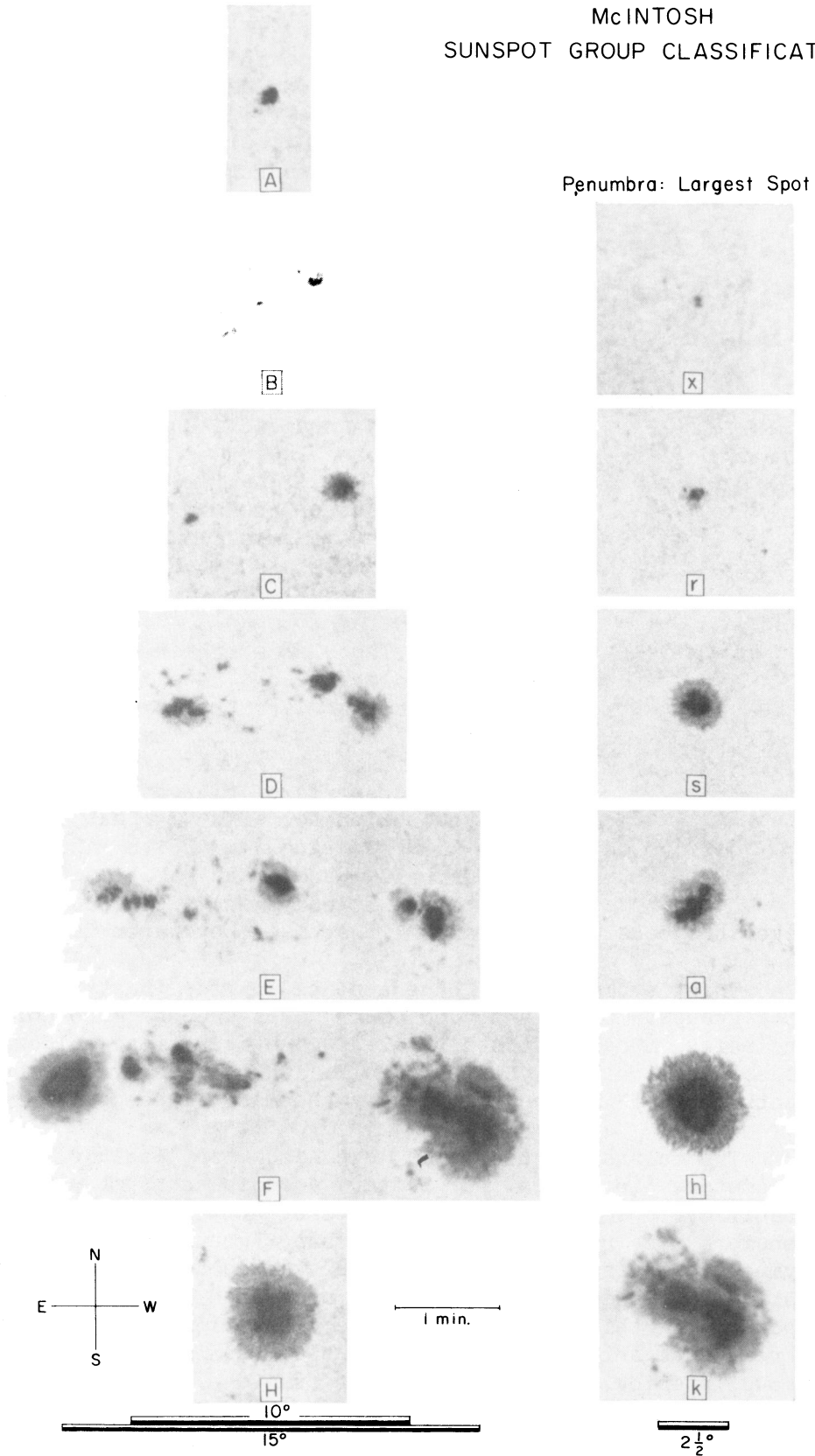


Figure 4

Sunspot Distribution

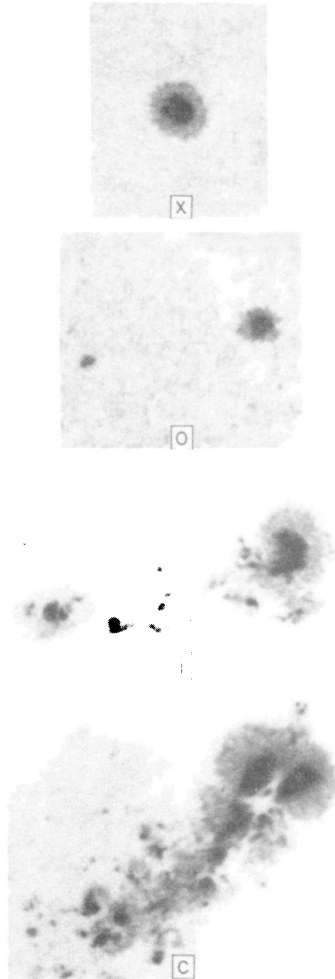


Figure 5. The third parameter of the McIntosh Revised Sunspot Classification.

fit neatly within the scheme of classification used for rapid communication of sunspot information (Figures 4 and 5; McIntosh, 1966, 1972d, 1977). This regularity among such variety encourages our search for physical understanding of sunspots.

The current sunspot classification system is a three-digit scheme composed of class - largest spot - distribution. Each parameter is a simple progression in scale and/or geometry, so the application of the classification is not as complex as it appears. The definitions of sunspot-group class were borrowed and revised from the Zurich classification (Waldmeier, 1947, 1955), which in turn, had borrowed from the ideas of Brunner and Cortie (1901). The revision imposed more quantitative definitions and removed subjective judgements of group complexity.

The classes are defined on the basis of unipolarity or bipolarity, the presence or absence of penumbra, whether penumbra occurs on just one or both ends of the group and the length of the group in heliographic degrees. The sequence of classes represents the order of evolution for the largest groups, beginning with A, growing to F and decaying to H and finally back to A. Many groups decay to H after reaching only C or D. Half of all groups

never progress beyond A and E; that is, they never develop penumbra.

This evolutionary progression applies best to isolated, bipolar spot groups. Closely-spaced spot groups undergo more complex evolution, including mergers that result in configurations seldom repeated. The classification reflects "normal" sunspot behavior, but the most "pathological" groups are usually the most flare-active.

The classification was expanded to include two additional parameters in order to better describe sunspot size, complexity and stability. The categories of largest spot are based on whether penumbra is rudimentary or mature, whether the spot is primarily symmetric or asymmetric and whether the spot exceeds $2 \frac{1}{2}$ degrees in latitudinal width. Symmetric spots are longer lived than asymmetric spots. The third parameter in the classification divides groups into three categories according to the relative crowding of spots in the group interior, borrowing somewhat from the classes of star clusters. Open groups have weak gradients between opposite magnetic-polarity spots; compact groups have strong spots with penumbra near the line of polarity inversion and, therefore, have the highest magnetic gradients.

The three classification parameters permit the definition of 60 distinct types of sunspot groups. This more detailed description of sunspots was adopted in 1969 by the International Ursigram and World Day Service for regular interchange of solar data worldwide. Correlations between flares and this new classification have been surprisingly good, leading to improvements in solar-prediction services (Kildahl, 1979).

The spot group in Figure 1 would be classified Dai. It produced an Importance 2 flare on the following day. Its trailing, or following, spot (east end of the group) is the largest in the group. Only 5% of spot groups have such follower dominance.

2.5 Sunspot Quantization

The illustrations for the sunspot classification (Figures 4 and 5) serve to show the range of sizes that occur among sunspots, and to suggest that some sizes are often repeated as if there were a quantifying process. Bumba (1965b) examined 97 leaders of recurrent spot groups and found that the spot areas occurred preferentially at quanta of 190, 310 and 570 millionths of the solar hemisphere. The Sacramento Peak films of smaller, shorter-lived groups suggests another frequent areal peak at 7 millionths, corresponding to a diameter of 5000 km and equivalent to the scale of the newly-reported mesogranulation (November et al, 1981). Spots smaller than this size seldom persist more than one day. Spots near this size often coalesce to form the larger umbrae. The quanta reported by Bumba may be thought of as multiples of the mesogranulation scale, and the first two of his values correspond to diameters of 2.7×10^4 km and 3.5×10^4 km respectively, which correspond to the average diameter of the facular network and the cells of the supergranulation network (Leighton et al, 1962).

The sunspots in Figure 6 were printed to identical scales and were selected as representing the maximum size for symmetric spots and cellular groupings of spots. The largest elements of even complex spot groups are identical in scale to the largest of the isolated, symmetric spots. The cellular arrangements of smaller spots, almost always occurring in the trailing portions of active regions, occur at a scale nearly identical to the large, symmetric spots, as if leader spots fill cells and trailing spots outline the cells. The size of these symmetric spots and cells, again, is curiously close to that of the network and supergranulation cells.

It is possible that the entire sunspot group is quantized in multiples of network cells, as it is common to observe chains of spot in arcs that partially outline cells of this scale [Bumba, 1965a (Figure 9)]. Penumbra on spots within these arcs is almost always directed away from the center of the cell, which is also away from the interior of the spot group. The exception to this rule in Figure 1 must be due to the presence of the nearby line of polarity inversion.

SUNSPOT QUANTIZATION

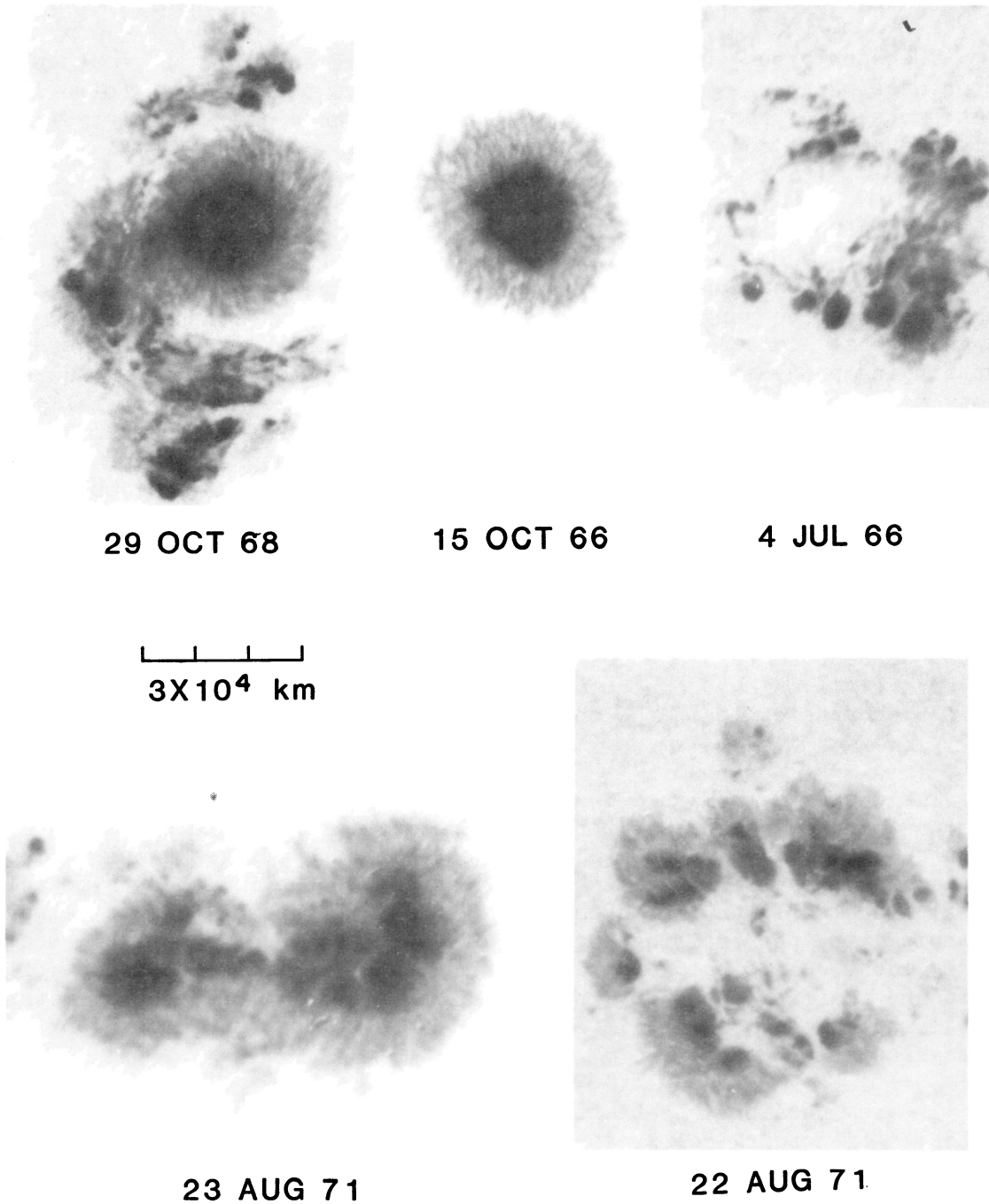


Figure 6. Examples of the largest leader sunspots, at left, occurred at a scale remarkably similar to the cells outlined by spots in the trailing portions of active regions (right). These circular features are identical in scale to the calcium network and supergranulation.

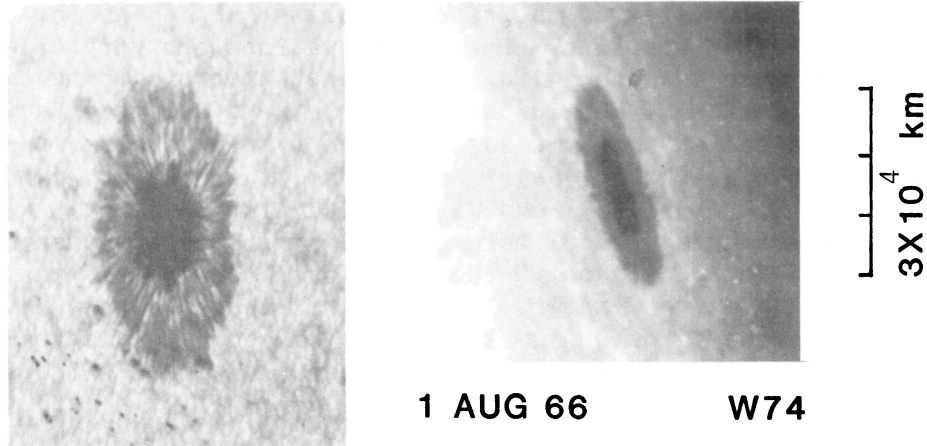


Figure 7. Examples of sunspots which did not show the classic Wilson effect. Left: Spot at W60 recorded by Stratoscope in 1959 (Danielson, 1961). Right: Spot recorded with the Sacramento Peak white-light patrol (Wilson and McIntosh, 1969). The Stratoscope photograph is about twice the scale of the patrol photograph.

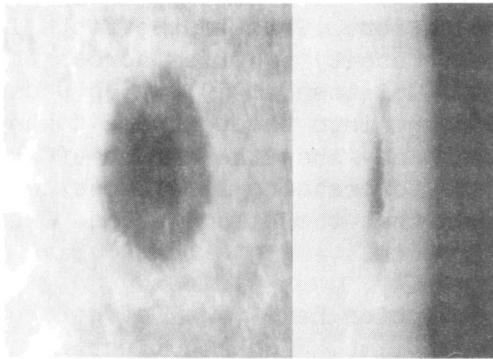


Figure 8. The classic Wilson effect displayed on 27 March 1963 in a spot only 5" arc from the solar limb. The photograph on 25 March was when located at W60. The spot measured 14,000 km in diameter. Photographed with the Sacramento Peak 12-inch (30 cm) horizontal telescope.

2.6 The Wilson Effect

The appearance of sunspots as depressions in the solar atmosphere was first observed in 1771 by both Schülen (Newcomb-Engelmann, 1948) and Wilson (1774), and more recently confirmed by Loughhead and Bray (1958), Chystiakov (1962) and Suzuki (1967). These measurements and descriptions imply that the photosphere occults the penumbra toward the center of the solar disk, leading to an increase in the ratio of limb-side to disk-side penumbral width as the spot approaches the solar limb. This is the classic Wilson effect. In addition, Loughhead and Bray reported that the center-limb spot variation includes an increase in diffuseness of the disk-side penumbra-photosphere boundary while the limb-side penumbra becomes sharper in boundary. The disk-side penumbra also darkens near the limb, leading to a more diffuse umbra-penumbra boundary as well. This photometric Wilson effect has been confirmed by Beckers and Schröter (1969) and Wilson and McIntosh (1969). Interpretation of these observations has provided basic information on emission and absorption coefficients in a spot and their variation with depth. Our understanding of the physics of sunspots, then, has a fundamental dependence on the correctness of the Wilson effect.

A critical review of the papers dealing with the classic Wilson effect reveals that many of the observations are very old (Chevalier, 1919), most were obtained with small instruments during poor seeing conditions (Baxter, 1960; Suzuki, 1967) or were reported in tabular form without indications of measuring errors and quality of observing conditions (Chystiakov, 1962). Only those of Bray and Loughhead (1958), Beckers and Schröter (1969) and Wilson and McIntosh (1969) were derived from rapid-cadence patrol telescopes that provided thousands of frames from which to choose only those exposed during superior seeing conditions; and yet their observations are of only three cases. Chystiakov's review of the literature prior to 1960 revealed that a substantial number of spots (20-25%) showed no Wilson effect or appeared to have a reverse effect (as if the spot were an elevated structure). The case studied by Wilson and McIntosh exhibited a Wilson effect at its east-limb passage but remained symmetric throughout its western disk passage, with no hint of a Wilson effect even at 20" arc from the limb (Figure 7).

Chystiakov (1962) and Wilson and McIntosh (1969) argue that many observations of the classic Wilson effect may have been products of intrinsic spot irregularities, obscuration by overlying faculae, poor seeing or problems with photographic reproduction. This last effect is to be expected in view of the asymmetric intensity profiles across the spot (Beckers and Schröter, 1969; Wilson and McIntosh, 1969) which would encourage photographic "bleeding" of the umbra into the penumbra toward the center of the disk. The spot in Figure 8 may show the Wilson effect for a combination of these reasons. It is for reasons like these that Waldmeier (1955) and Unsöld (1955) suggested that the Wilson effect was psychological rather than astronomical in origin.

Systematic aspects of sunspot group evolution have the effect of creating "false" Wilson effects at either limb. For a collection of sunspots selected for their "ideal" symmetry at central meridian, spots will be younger at east limb and older at west limb. Leader spots typically have extensions of penumbral area toward their follower spots early in their lifetimes. Spots that are beginning to decay do so by fragmenting and losing penumbra on their trailing boundary. The east-limb (young) spots will tend to have intrinsic asymmetries with penumbra extended toward the limb (toward follower spots) and spots at west limb (older spots) will tend to have intrinsic asymmetries with an absence of penumbra toward the center of the disk. This explanation certainly does not remove all of the purported cases of Wilson Effect, but it may sufficiently reduce the number of valid cases such that the classic Wilson effect should not be the foundation of a generalized theory of sunspots.

The careful photometric work by Beckers and Schröter (1969) and Wilson and McIntosh (1969) and the interpretation of white-light images by Loughhead and Bray (1958) establish that there is, indeed, a center-limb variation in the intensity profiles of spots; and, these observations are consistent with the spot umbra being depressed below the level of the surrounding penumbra and photosphere. However, it remains unexplained why some sunspots change their proportions of penumbral width as they approach the limb while others do not.

Additional observations are needed, and are readily available in the excellent white-light patrol films accumulated at the Sacramento Peak National Observatory since September, 1963 (McIntosh, 1964). Only a continuous, rapid-cadence patrol can accumulate complete disk transits of many examples so that the effects of spot asymmetry, spot evolution and seeing conditions can be analyzed to determine the true nature and frequency of the Wilson effect.

3. The Birth of Sunspot Groups

Most of the published data on the birth of sunspots applies to the formation of spots in groups already established (Loughhead and Bray, 1961) or come from observations made at 24-hour intervals (McIntosh, 1962). The critical observations of the earliest aspects of spot formation and how spots develop as a bipolar group can come only from a continuous, full-disk patrol such as that at Sacramento Peak. A review of the rapid-cadence patrol films permits tracing back in time from the time a new group is conspicuous to the time when it made its first appearance as a tiny sunspot pore, carefully noting the coordinates of the spots during the review process. Such a thorough review was made of all the films during the first two years of the patrol, examining 225 spot groups during a time of sunspot minimum (1963-1965). An additional hundred groups have been examined during periods of special interest since that time. Only 15 bipolar groups in this data set formed at such a time during the Sacramento Peak observing day, and with favorable observing conditions, to allow detection of the first sunspot formation. The leader sunspot formed first in ten out of the 15 groups. In only two cases could the time of formation of both leader and follower spots be determined reliably. Both of these cases were when the follower spot formed first. In one case the time between formation of follower and leader was seven hours; in the other case the time difference was about 26 hours.

3.1 Proton-Flare Sunspot Group of July, 1966

Figures 9, 10 and 11 illustrate the birth and early evolution of the sunspot group that was the subject of study by the international Proton Flare Project (Stickland, 1969). Although this group became very complex at maturity, its initial stages are typical of most sunspot groups. The first spot (A) of follower (positive=N) polarity, formed on the boundary of a network cell that is outlined with faculae [visible near the center of the disk in this example through photography in a near-ultraviolet line that records faculae over all the disk (Dezso, 1966)]. This is in agreement with studies of the formation of active regions with respect to the calcium network structures (Bumba and Howard, 1965a; Born, 1974) wherein both the initial plage brightenings and the first sunspot pores were observed to form on the cell boundaries and, more specifically, at the intersection of two or three adjoining cells. The initial spots of leader=S polarity formed a few hours later on the opposite side of the same network cell, so that the separation of the initial leader and follower spots is almost always identical at $2.0 - 3.0 \times 10^4$ km, corresponding closely to the typical cell size of either

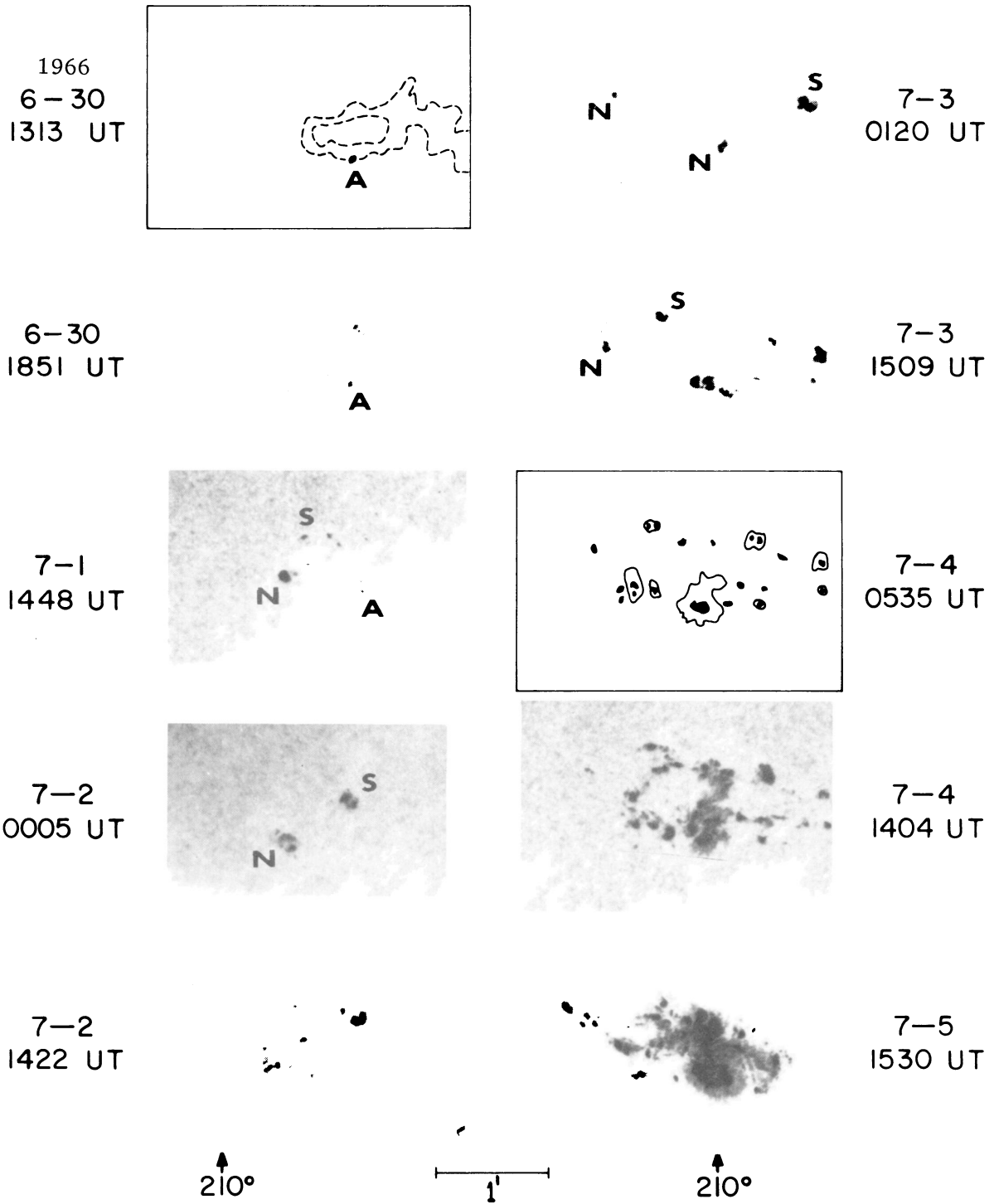


Figure 9. Birth and early evolution of a proton-flare sunspot group (McMath Region 8362) at N33, Carrington longitude= 210° (McIntosh, 1969a). North heliographic is up, west to the right, as in all figures in this paper. Drawings are based on photographs taken near 3750A by L. Dezso, Debrecen, Hungary. Photograph at 30/1851 U.T. from Aerospace Corporation. All others from Sacramento Peak white-light patrol.

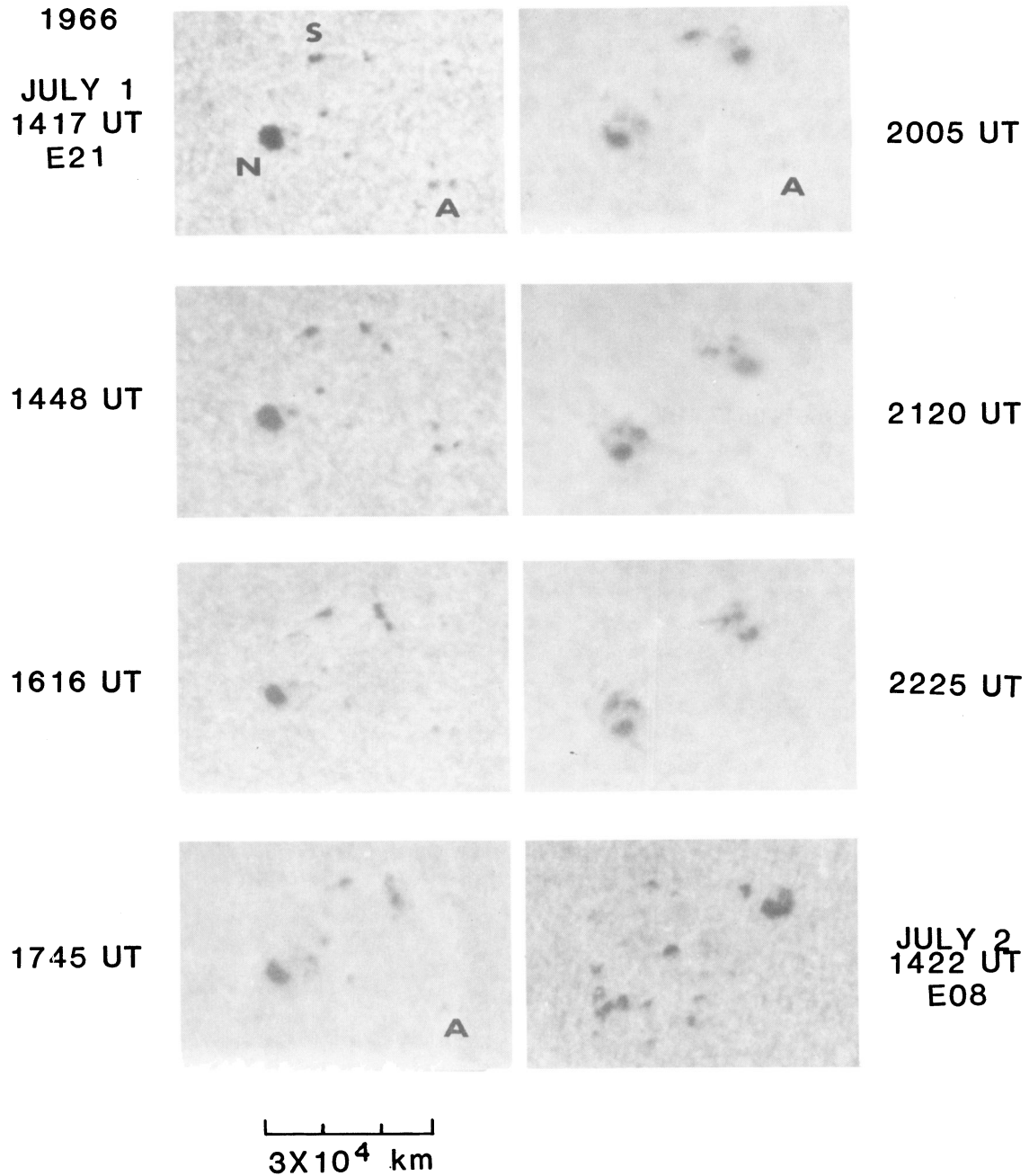


Figure 10. Detailed early evolution in McMath Region 8362 included: spot coalescence to form a larger leader spot, decrease in size and darkness of follower spot, counterclockwise vortical motion among the follower spots, a widening separation between the poles of the group and a slight clockwise rotation of the spot-group axis.

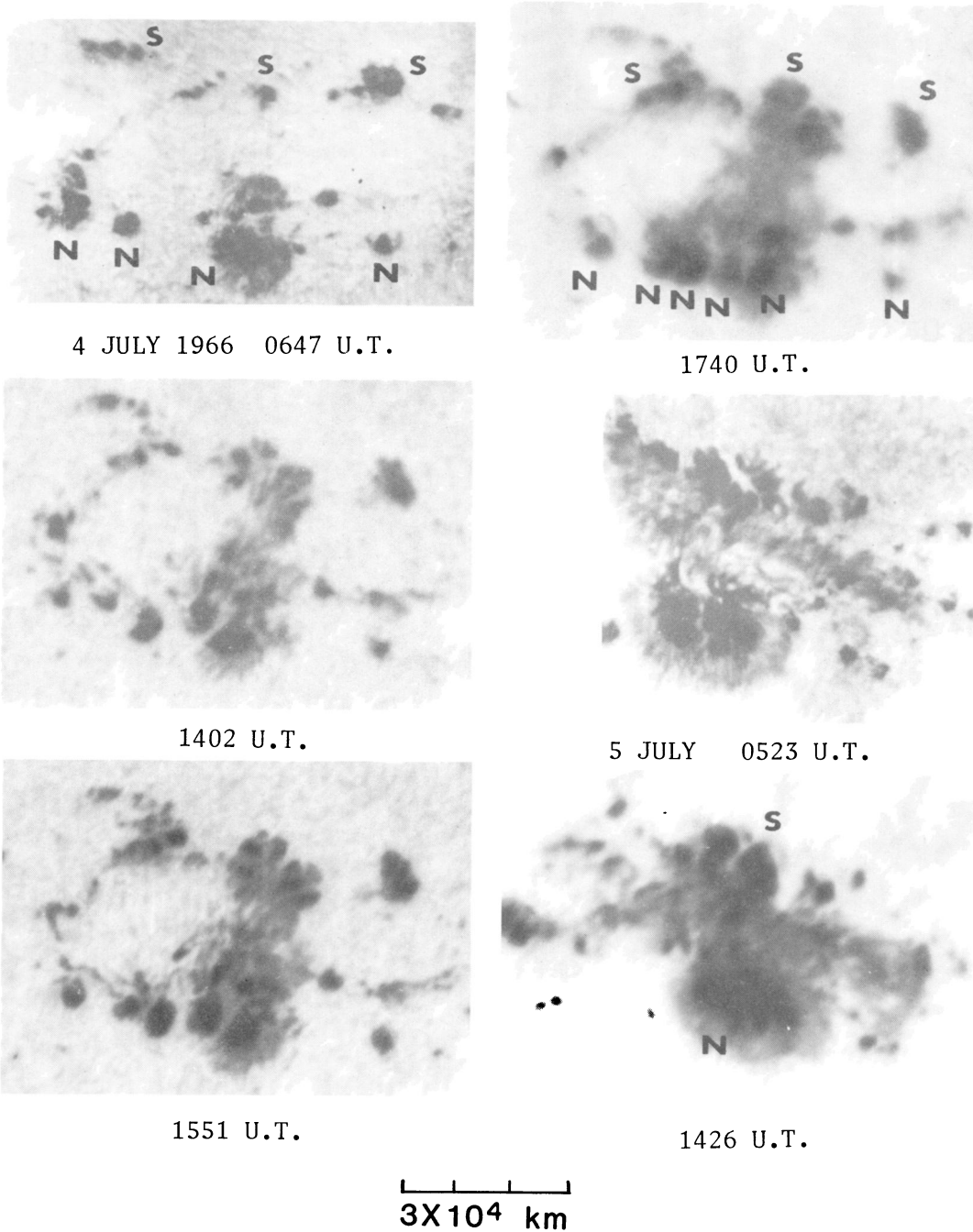


Figure 11. Rapid growth of spots and penumbra around and between network cells in McMath Region 8362. Coalescence of small spots forms major umbrae within the penumbra. Photographs at 4/0647 UT and 5/0523 UT courtesy of N. V. Steshenko, Crimean Astrophysical Observatory.

the network cells or the supergranulation. Never has a bipolar pair of spots been observed to form closer together than this, unless it emerged within a large and complex spot group. Certainly the formation of the initial spots of isolated spot groups obeys this pattern faithfully.

The initial spots in Figure 9 formed a bipolar group with its axis highly inclined to the solar equator, a not uncommon occurrence. Mature spot groups almost always have a small inclination to the solar equator, with the leader (western) spot slightly closer to the equator. We see in this example the process that shifts the group axis toward the normal inclination. The initial spots are replaced a day later with a new bipolar pair at a less-steep inclination, and yet a third bipolar pair emerged on the third day to complete the "rotation" of the group axis to a typical value of inclination. This rotation is not gradual, but occurs in sudden "jerks" (Weart and Zirin, 1969), which we see here must coincide with the emergence of each new bipolar set of spots. The direction of rotation was clockwise in this northern-hemisphere example.

The appearance of the second and third bipolar sets of spots, and the pattern of growth in Figure 11, may be interpreted as spots outlining successively adjoining network cells. The process of flux emergence seems to be guided by the supergranulation, while the event of sunspot group formation involves more than one adjoining supergranule cell.

The detailed view of the evolution on the second day (July 1) (Figure 10) illustrates the value of a rapid-cadence, continuous white-light patrol. Sufficient frames were exposed throughout the day that at least one frame each hour captured a moment of superior seeing conditions. The evolution during this 24-hour period included the decrease in size and darkness of the follower spot and three types of spot proper motions.

The first motion was the familiar diverging motion between leader and follower spots, which occurred in this example with an initial velocity of 500 m/s and later slowed to 250 m/s (McIntosh, 1969a). These values mimic the observed rate of expansion of calcium plage in new active regions (Eumba and Howard, 1965a).

The second motion was the coalescence of several small spots to form the larger leader and follower spots. These spots, near the top-center of the frames in Figure 10, moved along arcs that appeared to follow the boundary of the network cell on which the spots formed. The velocity of motions relative to one another ranged from 80 to 220 m/s.

The third type of spot motion was a counterclockwise vortical motion among the three spots in the cluster of follower spots, in the lower-left of the frames in Figure 10. The two smaller spots moved in an arc centered on the third, larger spot, traveling through an angle of 45 degrees between 2005 UT and 0005 UT. The linear velocity of translation was 300 m/s; the angular velocity was about 7×10^{-5} radians/s. The radius of the vortex was 4700 km. Again, the scale of mesogranulation appears in the activity of small sunspots.

The birth and movement of spots on 4 July continued to demonstrate an intimate relationship to network cells (Figure 11). During this period of most rapid spot growth, spots formed in long rows and arcs that outlined two adjoining cells of precisely the dimension of the calcium network, and the diameter of the cell to the left (east) was identical to the initial separation of spots in all three of the bipolar pairs that formed on the first, second and fourth days (Figure 9).

Spot coalescence occurred again with the movement of spots in the lower-left portion of the group toward the west to form the principal spot of follower (N) polarity. Even though these spots were much larger than those in Figure 10, the velocities along the border of the cell were even higher, ranging from 190 to 300 m/s.

The space between the coalescing spots brightened during the final hours before merging, as seen in the frame at 1740 U.T. This phenomenon has been observed regularly in subsequent examples of sunspot coalescence. These bright lanes between umbrae have usually been called "light bridges", but such a term has an improper connotation of an elevated structure that obscures underlying spot structure. It would appear instead that there is local heating of the photosphere caused by the merger process, as if there were resistance to the merger. This resistance is implied as well by the continued definition of the original spots by intensity maxima within the final large umbra, dividing that spot into components corresponding to the spots that merged to form it. This aspect is visible in the frames for 5 July in Figure 11, especially on the original negatives that more clearly show umbral details.

The bright streak of granulation entering the large spot in Figures 1 and 2 is more deserving of the term "light bridge". This structure also occurred in H-alpha images as a bright streak spanning the spot. The brightness contours of the large umbra to the north paralleled the "bridge" and only gradually darkened with distance from the "bridge".

3.2 McMath Region 8404 of July, 1966

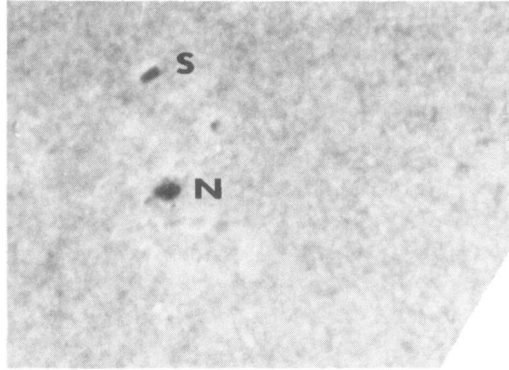
The birth and early evolution of the spot group in Figures 12 and 13 repeat some of the aspects of the proton-flare spot group, but in simpler form. The initial spots (Figure 12) formed with a north-south orientation and with the poles separated by 1.9×10^4 km, nearly identical in scale and orientation to the initial spots in Figure 9. The evolution during the first day was such to slightly rotate the group axis in a clockwise sense toward a more normal orientation. The completion of this change in orientation was accomplished by the emergence of two additional bipolar sets of spots on the second day, forming on what might be interpreted as the two network cells adjoining the cell outlined by the initial spots.

The initial spots in Region 8404 decayed rapidly on the third day (24 July) (Figure 13) while an arc of leader spots from the third bipolar set began to move westward. During the next three days a chain of at least eight spots performed a spectacular example of coalescence

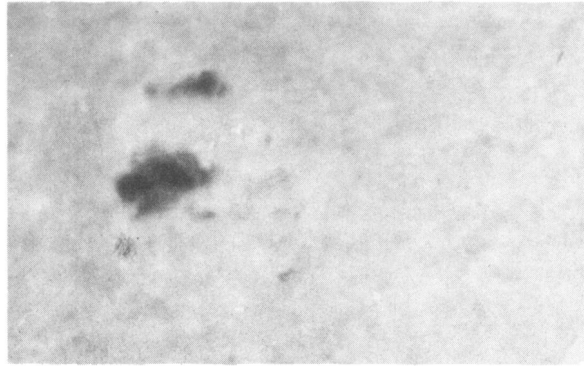
1966

22 JUL
1502 UT

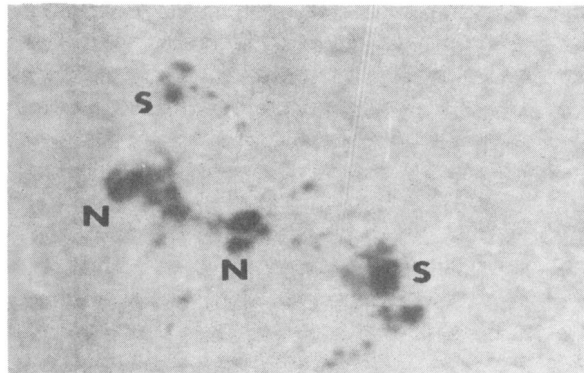
┌───┐
3X10⁴ km



22 JUL
2325 UT



23 JUL
1506 UT



23 JUL
2335 UT

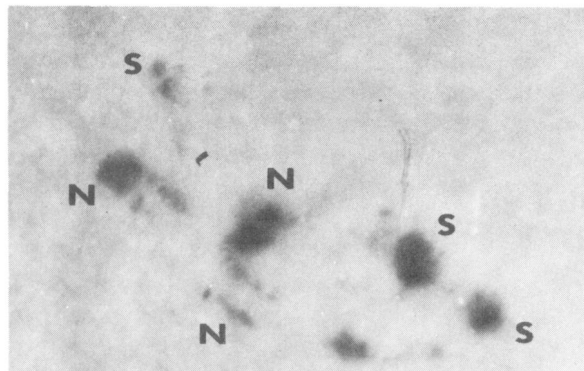


Figure 12. Birth and early evolution of spots in McMath Region 8404, N35, Carrington longitude=270°. Initial north-south bipole joined by second and third bipolar sets of spots by the end of the second day.

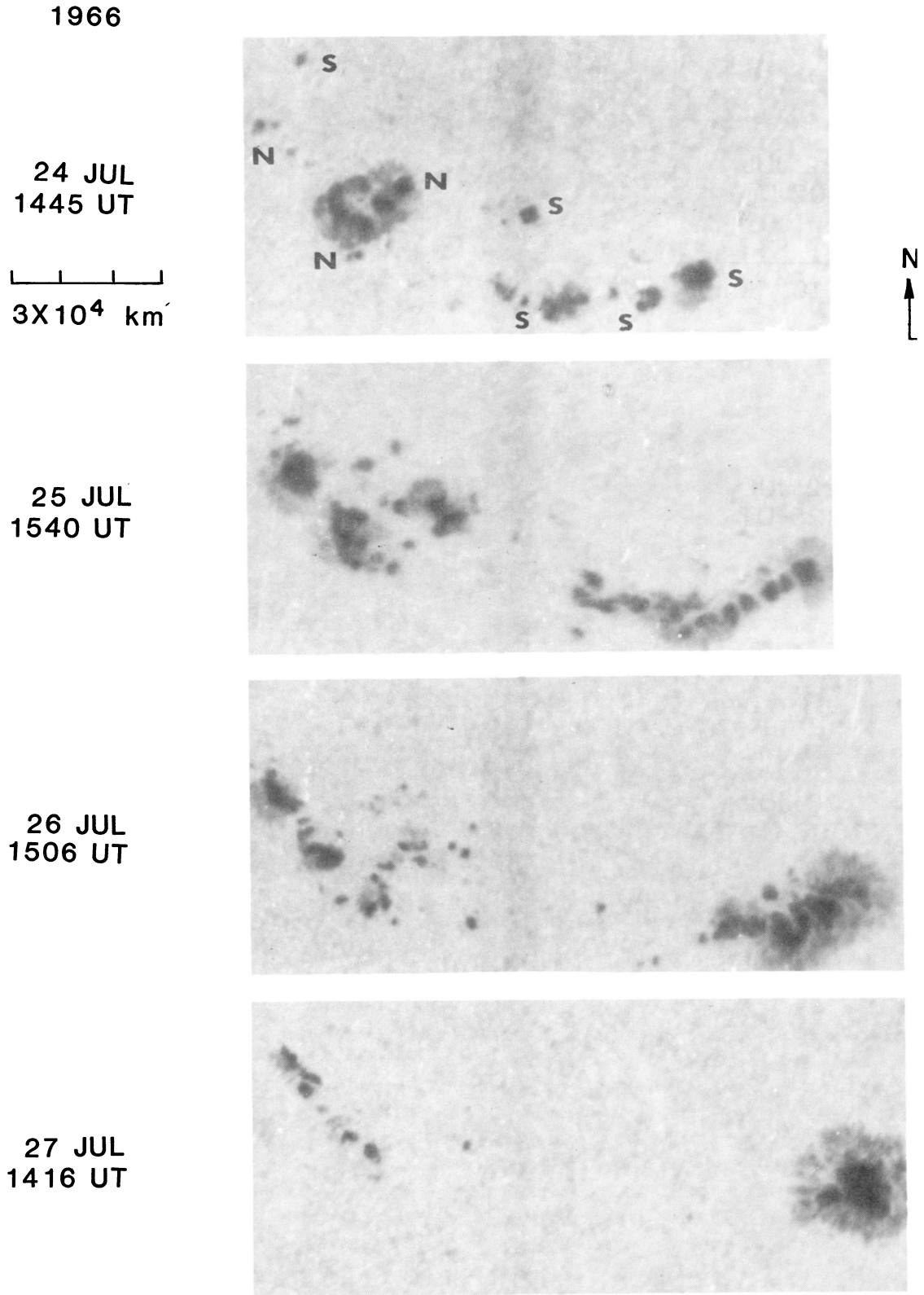
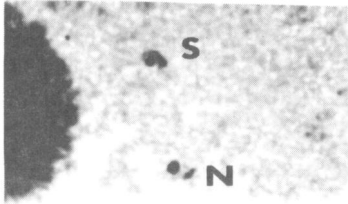
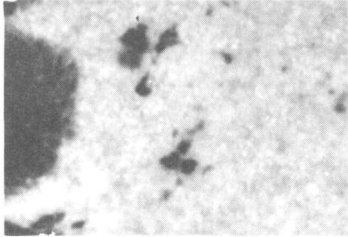


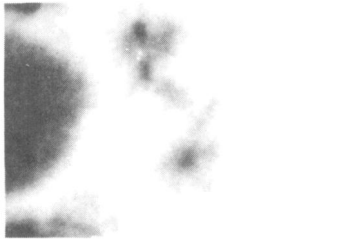
Figure 13. Coalescence of small spots formed the dominant, symmetric leader spot while the follower spots decayed and the entire group spread in longitude. McMath Region 8404, N35, Carrington longitude 270



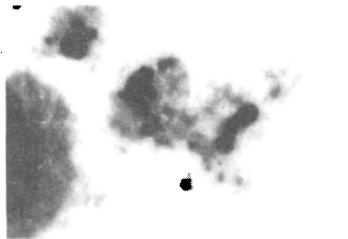
28 OCT
1504 UT



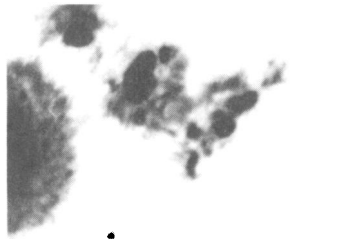
29 OCT
1637 UT



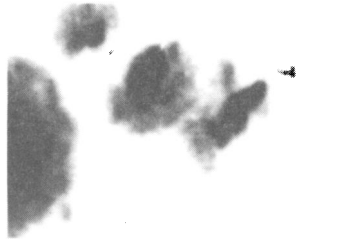
29 OCT
2222 UT



30 OCT
1454 UT



30 OCT
1706 UT



30 OCT
2144 UT



31 OCT
1530 UT

10"

Figure 14. Birth and counterclockwise rotation of axis of bipolar set of spots near the principal spot in McMath Region 9740, October, 1968 (McIntosh, 1970). A large proton-emitting flare occurred over these spots late on 30 October.

to form a symmetric leader spot. The velocity of the coalescing motion was 100 m/s. The follower spots underwent a steady decay without any appreciable motions other than the divergence from the leader spots. The rate of divergence was 140 m/s.

3.3 Proton-Flare Sunspot Group of October, 1968

The formation of a bipolar spot group in Figure 14 differs from the previous two examples in two respects: it occurred as a satellite group to a large and complex group, part of which is visible at the left edge of the photographs; and, it occurred in the sun's southern hemisphere. Again the spots formed with the group axis highly inclined to the solar equator, and rotated toward a normal inclination by the formation of new spots. The direction of rotation of the group axis was counterclockwise in this case, opposite to what was observed for the two cases in the northern hemisphere. These three examples suggest that the change in sunspot-group axis orientation follows a Coriolis law for vortical systems with outflow (C. Hyder, private communication). The supergranules have upward gas motions in their centers and weak flows directed from the center to the cell boundaries (Simon and Leighton, 1964).

A spot in the upper-left corner of the last four frames of Figure 14 had a proper motion of

120-200 m/s toward the newly-formed bipolar set of spots. This moving spot is also present above the main spot on 29 October, as seen in the upper-left portion of Figure 6. The movement of this spot during 28-30 October was in an arc centered on the main spot as if participating in a vortical motion with a clockwise sense; the penumbral filaments surrounding the main spot also had curvature appropriate for a clockwise vortex (McIntosh, 1970). Note that this sense of vorticity is opposite that in the small follower cluster in the northern-hemisphere example in Figure 10.

3.4 Origin of Longest-Lived Spot: McMath Region 8454

The observations of the birth and early development of McMath Region 8454 in August, 1966 were interrupted by poor weather, so it was not possible to determine whether it mimicked all the aspects of the previous examples. The importance of the observations in Figure 15 and 16 is in their record of the formation of the large leader sunspot that became the longest-lived spot in at least the past 18 years (137 days; 1.18×10^7 seconds; five complete solar rotations).

No spots were observed at this location during the latter half of 22 August, so it is assumed that the initial spots formed during the first 20 hours of 23 August. The spots formed into two parallel rows by 24 August, with the larger, leading row north of the followers and separated from them by the usual network-cell diameter. The granulation between the rows was streaked with dark lanes, presumably mapping the strong transverse magnetic fields connecting the spots of opposite polarity.

The row of leader umbrae were very similar to those of Region 8404 on 25 July (Figure 13), and so we assume they coalesced to form the symmetric spot observed 21 hours later. The velocity of spot proper motions in the coalescence must have been 325 m/s, three times that of the spots in Region 8404.

The formation of this long-lived spot is distinguished from the major spots in the previous two examples by the occurrence of rotation among the coalescing spots. The major axis of the non-circular umbra rotated through 210° during 24-28 August, with the motion over successive 24-hour periods of 70, 50, 35 and 55 degrees. The spot was too close to west limb by 29 August for further measurements. The sense of rotation was counterclockwise, consistent with the Coriolis vorticity expected for solar differential rotation in the northern hemisphere. The penumbral fine structure on 27 August also had curvature appropriate for counterclockwise vorticity. This sense of vorticity was also present in the chromospheric fibrils surrounding the penumbra.

The occurrence of rotation among spots is not limited to leader spots, although the shorter lifetimes of follower spots allows for few examples. The eastern-most follower spots in Figure 16 coalesced on 26 August to form a pair of closely-spaced umbrae by 27/0046 U.T. That pair had rotated some 45° counterclockwise by 27/1512 U.T. This example, and the follower spots of Figure 10, indicate that there is not a reversal of sense of vorticity between leader and follower spots.

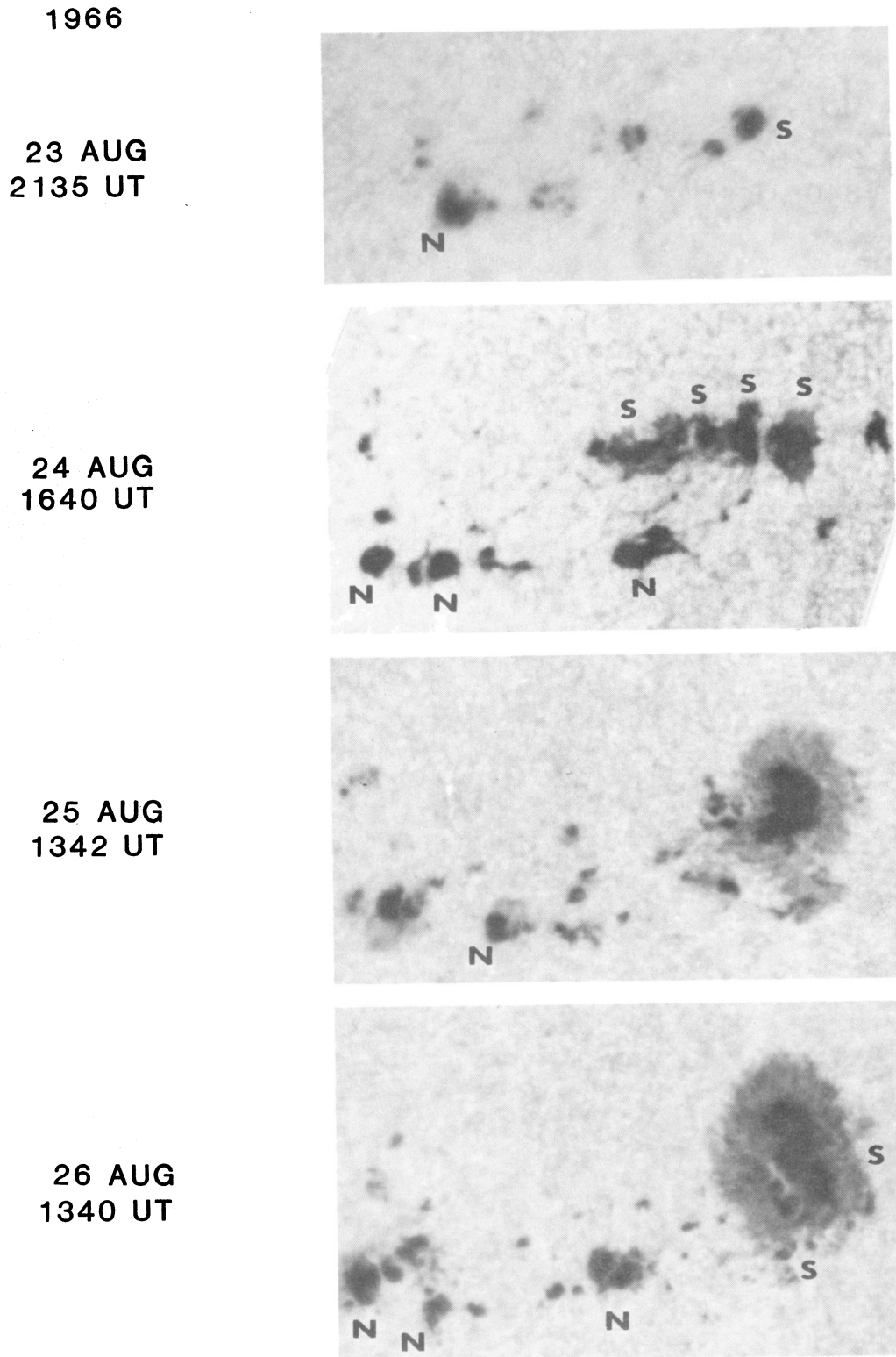


Figure 15. Birth and early evolution of sunspots in McMath Region 8454 included rapid coalescence of spots to form a large, symmetric leader spot and counterclockwise rotation of its umbral mass. This spot had a lifetime of 137 days, the longest between 1963 and 1981. Photograph on 24th from the solar site survey at Mt. Kobau, Canada.

1966

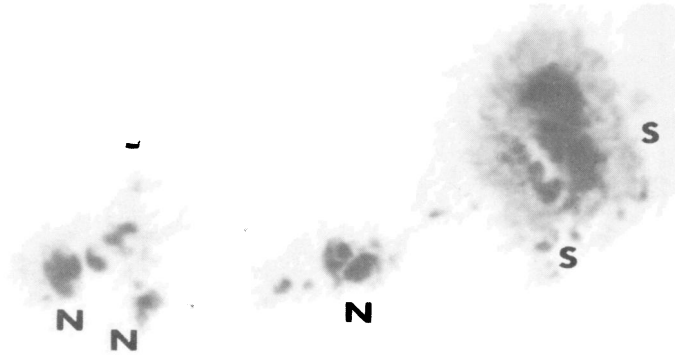
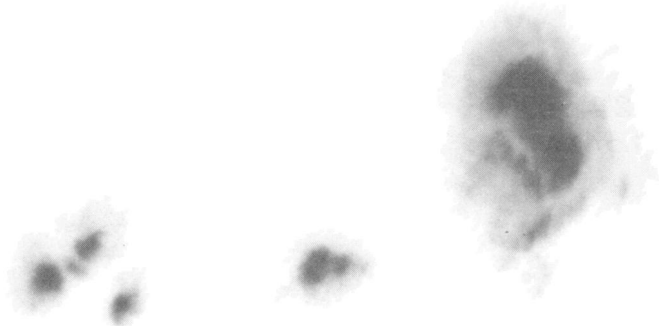
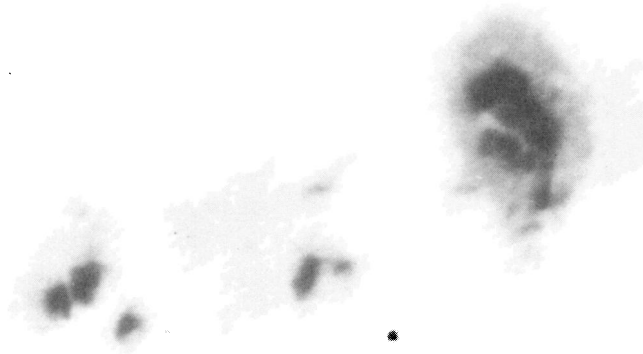
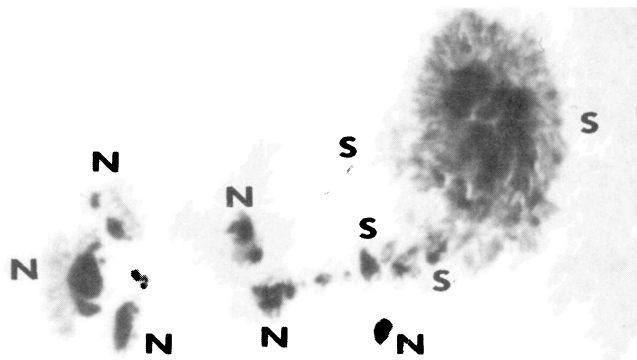
26 AUG
1340 UT26 AUG
1800 UT27 AUG
0046 UT27 AUG
1512 UT

Figure 16. Continued rotation of leader umbral mass in McMath Region 8454 is most evident by observing the changing orientation of the light bridge left(east) of the umbrae. Vortical fine structure in the penumbra on 27 August is additional evidence for counterclockwise rotation.

4. Spot Dynamics in Mature Groups

Sunspot proper motions are most extreme in the sunspot groups of proton-flare regions. The two examples illustrated in this section establish most clearly that spot rotation does occur, and that it follows the Coriolis law of clockwise in the southern and counterclockwise in the northern hemisphere. These groups also exhibited dramatic examples of spot coalescence, adding important evidence that coalescence systematically separates opposite-polarity spots, gathers together like-polarity spots and results in temporary brightening of the photosphere between spots during the final hours of coalescence.

4.1 McMath Region 8461 of August, 1966

The sunspot rotation so obvious in Figure 17 demonstrates that spot rotation does not imply stability, as it might have seemed in the case of McMath Region 8454 (Section 3.4). The onset of rotation simultaneous with important spot growth, and especially the acceleration of that rotation in the hours prior to a powerful flare, are evidence that vorticity accompanies the arrival of significant magnetic flux to the solar surface and that the motions are integral to the storage and release of flare energy. Sunspots represent magnetic energy available for the production of flares, but the expectation that the occurrence of large flares should reduce sunspot area is denied by this example. The spot area more than doubled in the eight hours following the great flare on 28 August. The flare energy must, therefore, come more from the relative motions of the sunspots than from their magnetic flux.

The three large spots early in the sequence in Figure 17 were all of negative (leader) polarity. Spot A was stationary 25-26 August and then accelerated during the 36 hours prior to the great flare on 28 August, reaching a velocity of 200 m/s as it was ejected from the penumbra toward the right (west). The spot then slowed to 50 m/s by the 31st. The initial bright points in the flare occurred near this spot (Zirin and Lackner, 1969) at the position where it would separate from the penumbral area only minutes after the flare maximum.

The other two spots in the penumbra rotated about a common center, completing a 180-degree turn between 25 and 29 August, during which time their areas increased fourfold. Simultaneous with their rotation and growth new spots of opposite polarity developed on the southeast and southern edge of the group, forming a configuration with the rotating spots on 27 August that resembled the cellular features on 4 July 1966 (Figure 11), including the position of leaders north of followers. The distance from the rotating spots to the opposite-polarity spots was 2.5×10^4 km, again suggesting the influence of the supergranulation, and permitting the interpretation that the rotating cluster of leader spots lay at the intersection of adjoining network cells. Spot A appears to have left this intersection to move along the border of the adjoining cell.

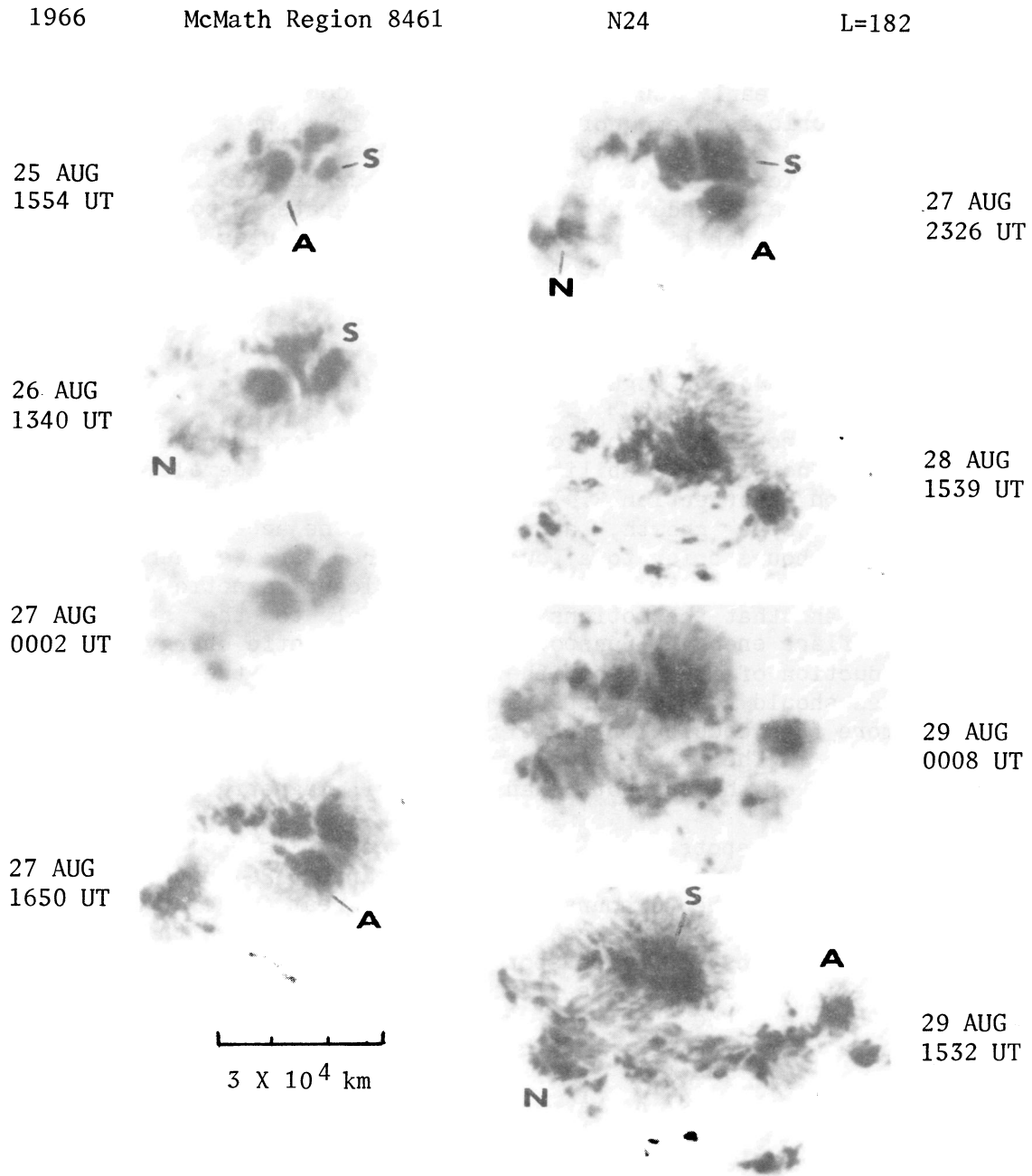


Figure 17. Counterclockwise spot rotation accelerated in this group more than a day before the great proton flare of 28 August 1966. Flare maximum in hard X-rays was at 1529 U.T. and H-alpha maximum at 1533 U.T. (Zirin and Lackner, 1969). Rapid growth of spots and penumbra began by the time the flare had subsided. The rapidly-moving spot separated from the main penumbral area less than an hour after flare maximum.

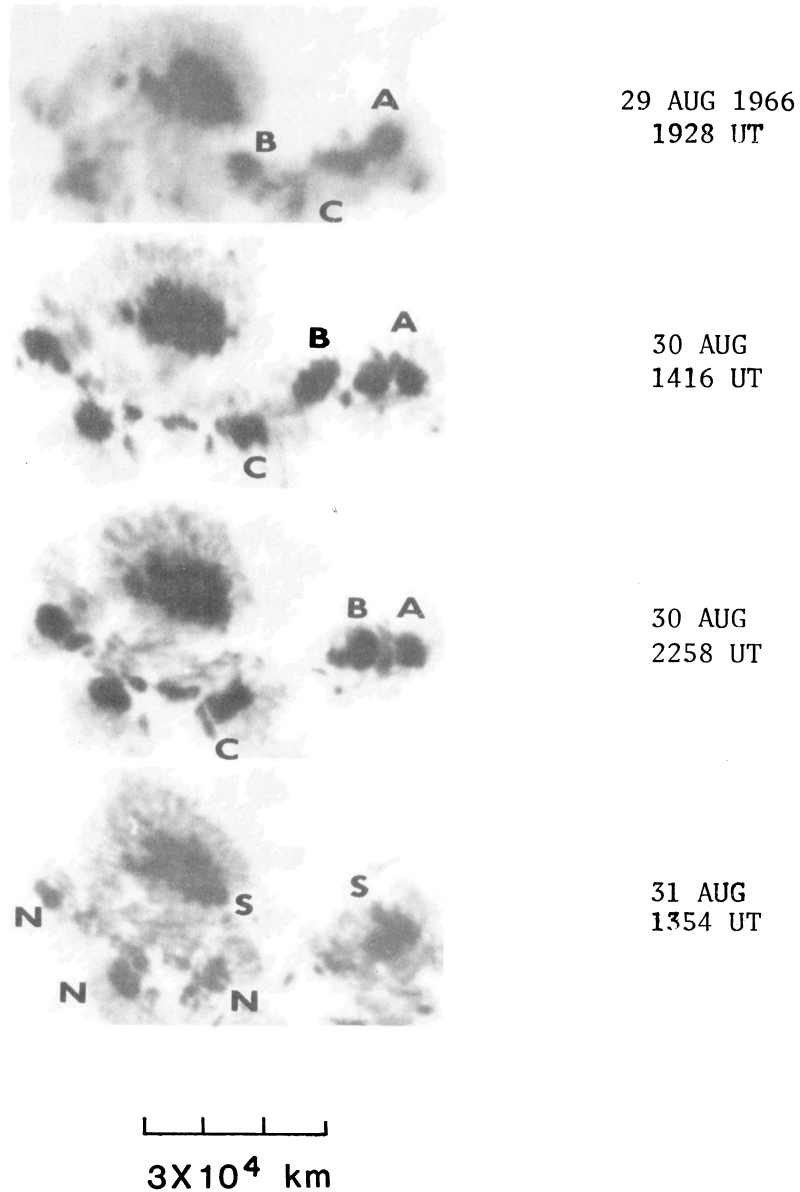


Figure 18. The evolution of McMath Region 8461 continued from Figure 17. Spots that coalesced to the right (west) were of leader polarity; spots that converged to the left were of follower polarity.

The spots of follower polarity did not participate in the rotary motions of the leader spots, as if the line of polarity inversion (neutral line) separated different regimes of mass motions in the photosphere. The shear introduced in the area of the neutral line by the relative motions of the spots was manifest in the system of loop prominences observed in H-alpha above this line following the great flare on 28 August (Nolan et al, 1970). The bright strands of loops lay at an angle to the neutral line, in a direction consistent with the relative spot motions.

The three leader spots maintained their separate identities throughout the rotation, although the bright lanes separating them became so narrow as to become invisible in Figure 17 late in the sequence. This resistance to complete merging between even spots of the same polarity is common to all the examples in this review.

Periods of spot growth often occur intermittently throughout a spot group's lifetime, especially in those that are large and flare-active. Figure 18 continues the evolutionary history of the proton-flare group of Figure 17, showing spot coalescence and divergence during a phase of spot growth late in the group's history. The east-west "arm" of penumbra on 29 August ran along the line of polarity inversion so that umbrae on either side of it were of opposite polarity. The spot A on the western end of the "arm" was the rapidly-moving spot ejected during the period of counterclockwise rotation in Figure 17. The "arm" appears to map the boundary between network cells.

Two new spots formed on the northern edge of the "arm" late on the 29th and began moving toward the western spot. These three coalesced by the 31st, with the middle spot becoming highly compressed between spots A and B.

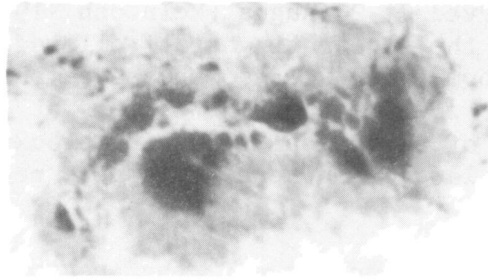
Spots of opposite polarity formed on the southern boundary of the "arm" on the 30th and moved in the opposite direction, approaching, but not achieving, coalescence with existing spots of follower polarity. Again we see motions change direction across the neutral line, with important shear in the magnetic field as a consequence. This shear is surely an important ingredient in the flare production of this region.

4.2 Spot Interpenetration

The giant spot group of August, 1971 (McMath Region 11482) also exhibited large, vortical, proper motions among its principal spots, including what must be one of the most singular examples of spot coalescence. During the period illustrated in Figure 19, the large umbra in the center of the leader penumbra moved at 350 m/s westward and through the largest umbral mass! The excellent photograph obtained by Bumba and Sykora (1972) shows a brilliant ring that formed around the moving spot, again suggesting that photospheric heating accompanies the process of merging spots of like polarity. The ring was also visible in the H-alpha line. All umbrae within this penumbra were of leader (positive) polarity.

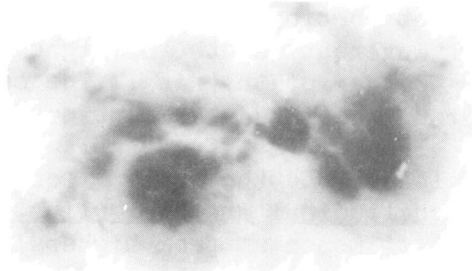
A series of very large, dark surges emanated from the southwest perimeter of the leader penumbra on the day of spot penetration, with the vector of surge motion oriented directly away from the site of penetration (McIntosh, 1972e; Zirin and Lazareff, 1973).

The penetrating spot formed at the southeast corner of the penumbra on 19 August and moved in an arc around the top of the large umbra in the eastern half of the penumbra. The large spot through which this spot passed was observed first on 18 August located directly north of the large umbra in the eastern half of the penumbra, and it also moved

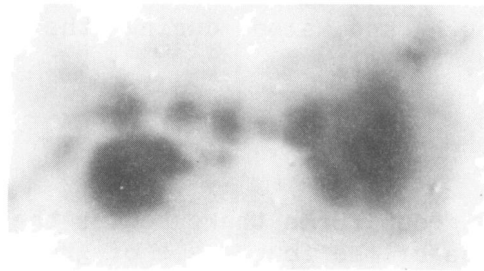


SUNSPOT INTERPENETRATION

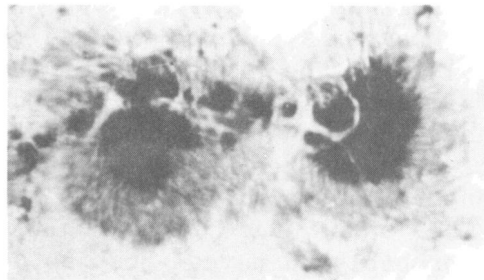
21 AUG 71
1345 UT



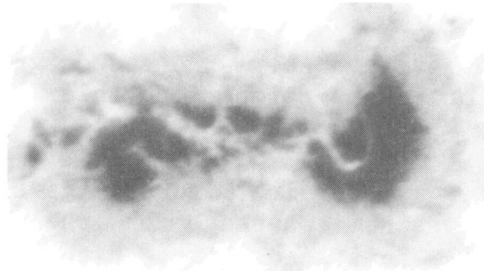
21 AUG
1745 UT




21 AUG
2310 UT



22 AUG
0535 UT
ONDREJOV, CZECH.



22 AUG
1513 UT


3 X 10⁴ km



23 AUG
1435 UT

FIGURE 19

in an arc to its position observed 21-23 August (McIntosh, 1972e). Note that the chain of spots left of the penetrating spot followed along the same path and by 23 August formed a row of compressed umbrae directly behind the penetrating spot. It would appear, then, that the large eastern umbra was the center of a vortical flow with clockwise sense, and this flow persisted in this large, southern-hemisphere region for at least six days.

The spots in the following portion of the region formed a configuration of a ring (lower right portion of Figure 6). The motions among these spots had a pattern of movement south (to higher latitude) then westward along the lower border of the ring. The movements of spots in leader and follower together depicted a continuous line of movement in the shape of the letter "S" from latitudes above the follower to latitudes between the leader and the solar equator. The line of motion paralleled the longitudinal neutral line. The spot motions then suggest the transport of momentum from higher latitudes to lower latitudes. The most rapid sunspot motions occurred close to the time of greatest sunspot growth and also during the time of most frequent flare activity (McIntosh, 1972e).

5. Sunspot Dissolution

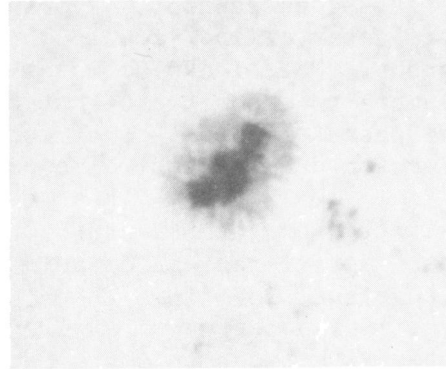
Almost as soon as a sunspot group forms there begins the process of sunspot dissolution. In every one of the four cases of group formation presented in this review the initial spots of each group survived only for a day or two, even though the group may have endured for weeks. Small spots labelled A on 1 July 1966 at the right edge of the frames in Figure 10 were the remnants of the initial bipolar set of spots for this group. They gradually decreased in size and darkness until they were only pores by 2120 UT. The spots did not break apart and the fragments scatter, as might be expected from prevailing ideas on dispersion of active-region magnetic fields. The spots, if they could have been observed at high resolution, probably decayed just as described by Bumba (1965a), with granulation gradually encroaching on the edges of the spot, and reducing its area by increments of a granule area. The size and intensity of the umbral granules probably increased gradually as well. Additional high-resolution observations of this process are needed.

The decay of the follower spots in Figure 13 gives a similar picture for larger spots. There was no tendency for the cluster of spots to separate as if being scattered. Instead they simply lost penumbra and decreased in size and intensity without appreciable motions.

The formation of large leader sunspots occurs through the coalescing of several smaller umbrae, but the formation of the smaller spots in a group occurs simply as a rapid darkening of intergranular spaces, followed by the enlargement of the spot as surrounding granules darken and become added to the edge of the spot, which is a reverse of the dissolution process for small spots. So we might expect that the dissolution of large leader spots might also differ from the dissolution

1966

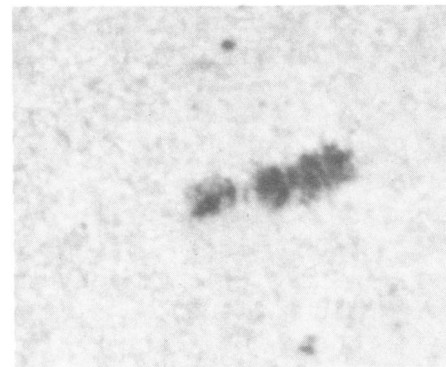
8 OCT
1408 UT



9 OCT
1526 UT



10 OCT
1455 UT



11 OCT
1710 UT

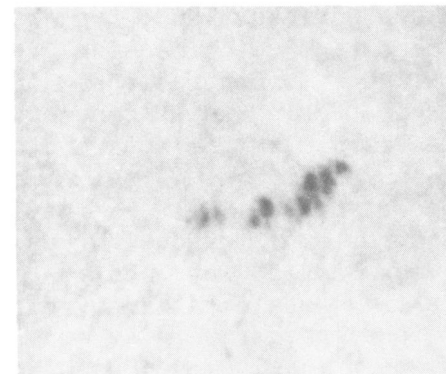


Figure 20. Dissolution of the leader spot of McMath Region 8530

of the smaller spots of a group. Figure 20 presents one example of the final days of a leader sunspot. A fragmentation process is observed, followed by a decay of individual umbrae much like that for other small spots of a group. The clockwise rotation of the major axis of the cluster of umbra is opposite to that observed during spot formation for northern-hemisphere leaders; but, there are insufficient examples studied so far for drawing general conclusions.

White-light observations cannot record evidence for the final stages of an active-region dissolution, as the magnetic fields and chromospheric plage sometimes persist for weeks and months after sunspots have disappeared. While no systematic study has been performed, observing experience has revealed that as a large sunspot fragments and decreases in area and darkness, chromospheric plage gradually forms surrounding and over the spot as if the magnetic fields from the spot were being replaced by fields represented by the plage. Inasmuch as these areas of plage gradually spread over more of the solar surface as the active region ages, one could conclude that the magnetic fields were being scattered. The difficulty with that conclusion is that no dispersive proper motions are observed for individual plage elements, just as there are no dispersive proper motions among decaying sunspots. The spreading and weakening of the active-region plage and magnetic fields seems to occur by the formation of new plage elements more distant from the center of the region as elements within the center fade and disappear. There is a similarity between this "scattering" process and the rotation of sunspot group axes by the successive formation of new spots and disappearance of old spots. There is a wave phenomenon moving the location of magnetic flux emergence.

6. The Large-Scale Connection

Observers as early as Carrington(1863) suspected that sunspot groups did not appear at random, and yet large-scale order to solar activity often eludes statistical studies(Howard and Edberg, 1973). The spatial distribution of solar activity becomes less random when statistics are based on larger events and active regions, leading several authors to declare the existence of "active longitudes"(Švestka, 1968; Bumba and Obridko, 1969; Švestka and Simon, 1969; Warwick, 1970). Studies of the extensive collection of Mt. Wilson synoptic magnetograms have emphasized concentrations of active regions into activity complexes, their persistence for numerous consecutive solar rotations, and their importance as the apparent source of giant regular structures in the large-scale patterns of solar magnetic fields(Bumba and Howard, 1965b; Bumba et al, 1968). The suggestion was made that the giant regular structures could be due to very large scale convective elements. A tendency was noted for active regions to form on pre-existing boundaries of large-scale patterns, but the low resolution of the Mt. Wilson data prevented definite conclusions. Ambrož et al(1971) ended their study with an admission of an empassé:

"It is not now possible to say whether the characteristic enormous pattern which develops is a natural consequence of the normal redistribution on the solar surface of the magnetic fields of active regions which are formed in more or less regular patterns across the

Sun, or whether the large-scale weak pattern which we observe is a direct reflection of large-scale subsurface fields which themselves produce active regions at the surface in a preferential pattern."

Limitations to the above studies of large-scale solar features were caused by variability of data over the long periods studied, low resolution and high signal-to-noise ratios. Averaging over two or more solar rotations was necessary in order to clarify the boundaries to large-scale structures. Separate displays of the two polarities were created in order to further reduce confusion in studying the patterns. A sense of connection between active regions and the giant regular patterns was perceived, but communication of the alleged connections was difficult and largely unconvincing.

A more convincing case for large-scale processes causing the emergence of active regions in preferred locations comes from the long term study of large-scale velocity fields (Howard and LaBonte, 1980). A display of velocity data over a 12-year period revealed long-lived, large-scale currents of both faster and slower velocities compared to the mean solar rotation. The boundaries between fast and slow streams define zones of excess large-scale shear. Maxima in the magnetic flux distribution map close to these zones of shear. Howard and LaBonte conclude that the solar magnetic (sunspot) cycle is not a result of random, drifting surface fields, but instead is the product of resonant, large-scale mass motions beneath the solar surface.

The large-scale velocity data had to be averaged extensively to improve signal-to-noise problems, preventing study of relationships to specific centers of activity. No comparison has been made with the giant regular structures revealed in the Mt. Wilson magnetograms.

6.1 H-alpha Synoptic Charts

A new perspective on large-scale solar magnetic fields and their relationships with solar active regions is provided by the H-alpha synoptic charts that are now available from 1964 to the present (McIntosh, 1979). Lines of polarity reversal inferred from systems of structures visible in H-alpha patrol filtergrams map the neutral lines in the radial component of solar magnetic fields, thereby revealing details to the boundaries of large-scale magnetic patterns that are not recorded in magnetograms sensitive to the longitudinal (line-of-sight) component of the fields. Figure 21 compares a recent H-alpha synoptic chart with the corresponding chart of magnetic fields measured at the Kitt Peak National Observatory. Negative polarity has been shaded uniformly black on the H-alpha chart to facilitate the comparison, and to emphasize the patterns of giant regular structures. The patterns of measured fields agree perfectly with the patterns outlined in H-alpha, with the exception that the boundaries to specific structures are defined with more precision on the H-alpha charts. In addition, the H-alpha charts define patterns to at least 70 degrees latitude, while the magnetograms display only diffuse features beyond 40 degrees latitude. The two charts are complementary. Field strength is a necessary datum in magnetic field studies, but the

**H α SYNOPTIC CHART
CARRINGTON ROTATION 1695 (PRELIMINARY)**

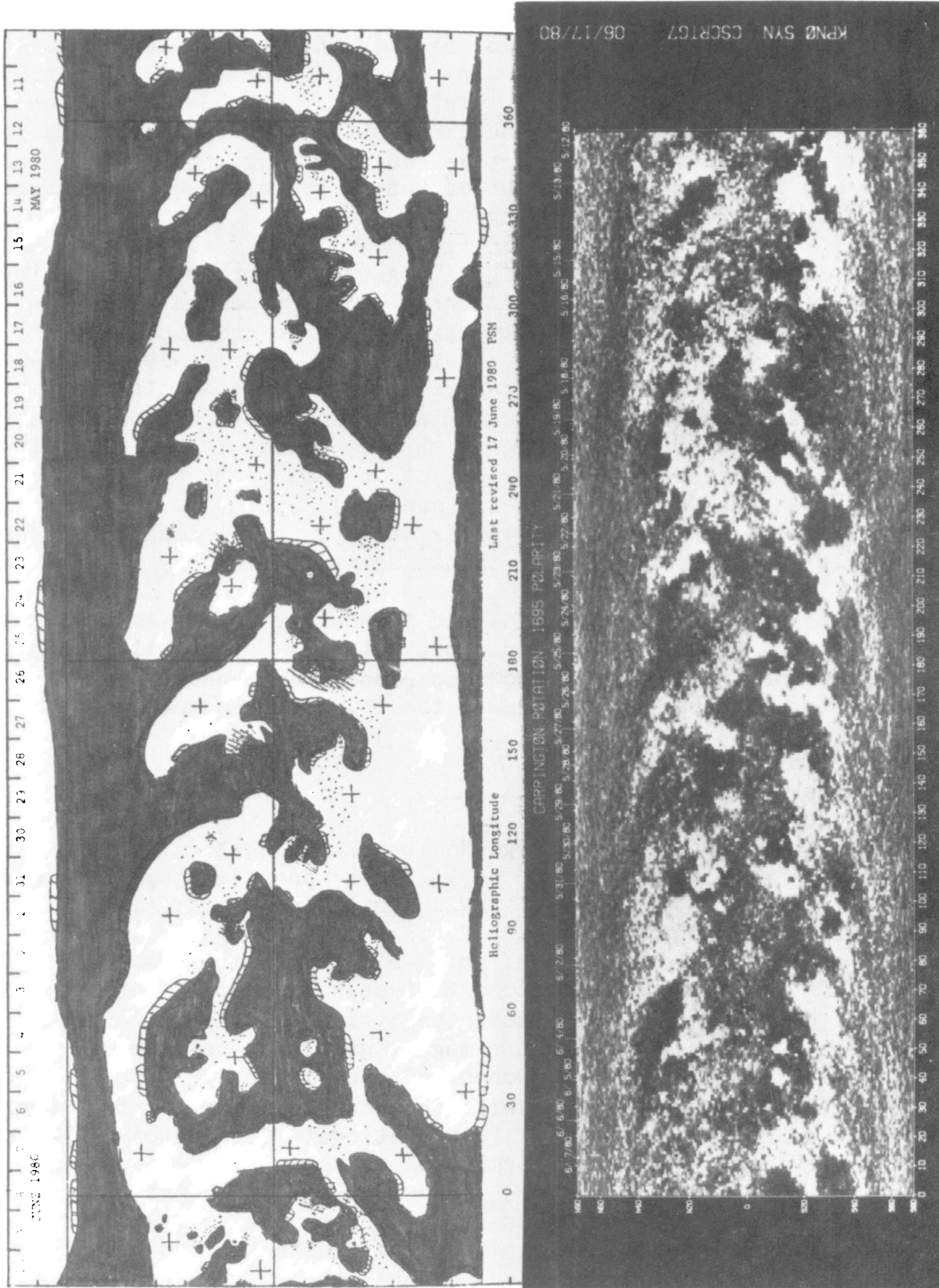


Figure 21

H-alpha neutral line is superior in defining boundaries to the large-scale structures.

The high resolution of large-scale boundaries permits a more instantaneous view of the dynamics of large-scale structures and a study of relationships to specific centers of activity. The H-alpha synoptic charts record active regions and large sunspots, and establish that the lines of polarity inversion within active regions usually connect directly to large-scale neutral lines bounding large-scale structures adjacent to the active regions. The atlas of H-alpha synoptic charts for Solar Cycle 20 (McIntosh, 1979) included notes on solar activity and apparent relationships to the large-scale aspects of solar activity, with many suggestions that large-scale processes originate and control the evolution of active regions (Section 6.5)

6.2 Large-Scale Patterns by Latitude Zones

The evolution of individual elements in the patterns of large-scale magnetic fields are followed more easily by dividing synoptic charts into limited zones of latitude and assembling time series showing single zones (Bumba and Howard, 1969; Ambrož et al, 1971; Bumba, 1976). The initial studies of the atlas of H-alpha synoptic charts divided the charts into three broad zones centered on the equator and latitude 40 in each hemisphere. These zones for the two-and-one-half years near the end of Solar Cycle 20 are presented in Figure 22, with negative polarity shaded black. The period of Skylab observations occurred near the end of this interval (Rotations 1601-1610).

The long lifetimes and large scales of persistent features are immediately apparent. This display confirms the existence of "rows" and "streams" discovered in the Mt. Wilson data (Ambrož et al, 1971), but with improved resolution and continuity. The detail at high latitudes far exceeds the displays constructed from magnetograms.

The gradual evolution of even the smaller elements in the patterns proves that the emergence of large sunspot groups does not introduce discontinuities in the large-scale evolution. Instead, sunspots form preferentially near the borders of large-scale structures on, or near, neutral lines with polarity arrangement appropriate for the Hale law for their hemisphere. The spots often occur at such times to assist the long-term evolutionary trends of large-scale features.

In exceptional cases, an especially strong active region will be followed by the gradual emergence of a major new element in the large-scale patterns. The largest and most-persistent dark structure in the equatorial zone during 1973 is just such an example. This giant cell of negative polarity was conspicuous during the Skylab observing period as The Fish (frontispiece, McIntosh, 1979). The diagonal formed by this feature can be traced back to Rotation 1590 and longitude 10, coinciding in time and position with the great proton-flare region of August, 1972. Magnetic-field strengths waxed and waned within this structure, accounting for the imperfect view of it provided by magnetograms. Outstanding active regions emerged repeatedly on the

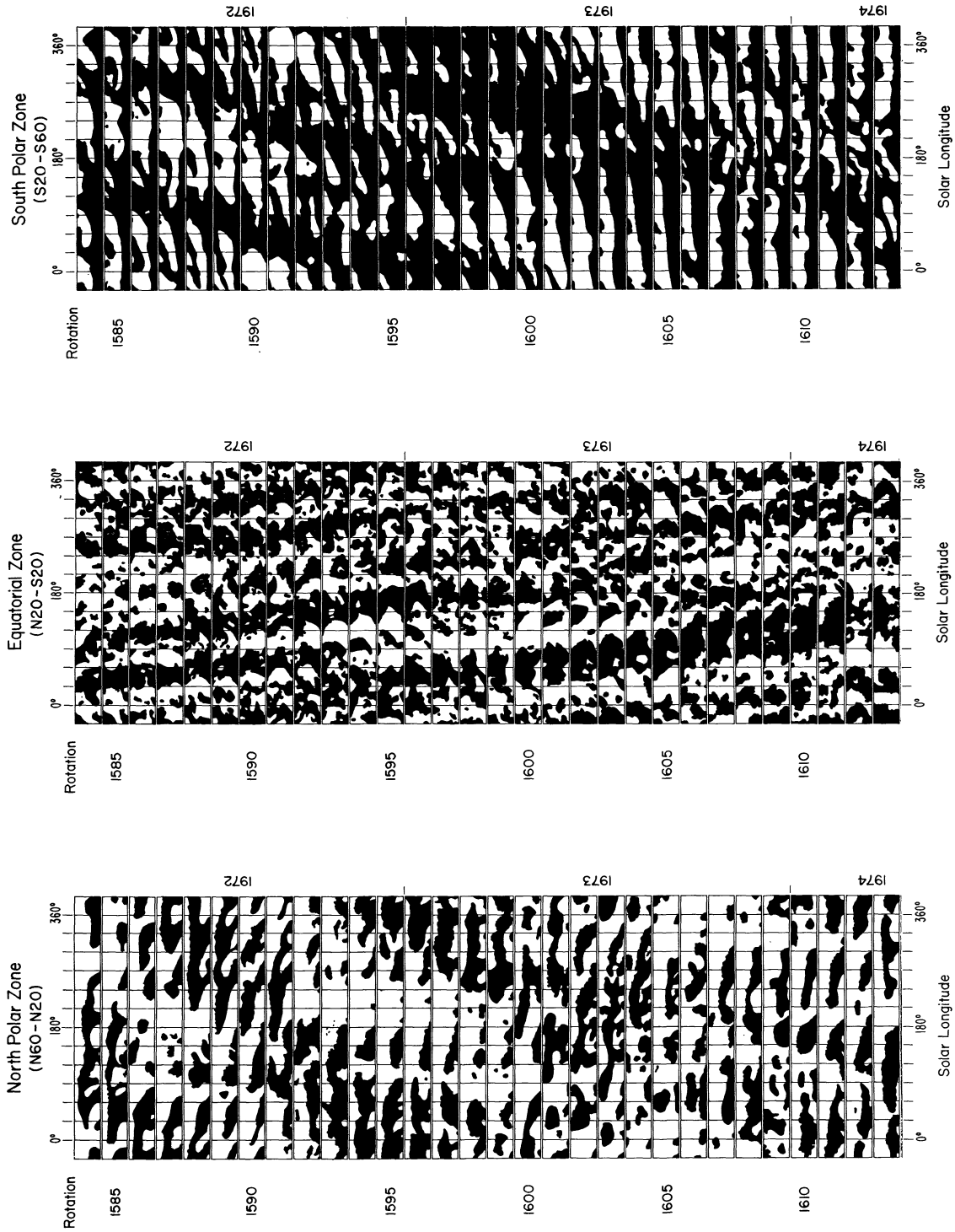


Figure 22

leading boundary of this structure and in the southern hemisphere, where positive polarity was leading in sunspot groups. It seems more than coincidental that this series of active regions culminated in another great proton-flare region in July, 1974. By that time the giant regular structure had drifted 180 degrees in Carrington longitude. This one example of activity involved in a particular giant cellular feature is substantial evidence that large-scale magnetic patterns are not consequences of sunspot formation, but instead are manifestations of the large-scale process (most likely convection) that generates sunspot groups.

The long-lived magnetic patterns drift westward (with respect to Carrington longitudes) in the equatorial zone, and eastward in the polar zones, reflecting differential rotation. Individual large-scale structures occur with persistent motions significantly above or below the average rate of rotation, leading to convergence and divergence among the large-scale structures. This phenomenon was particularly clear during the period of Skylab observations (1973) when patterns converged toward the middle half of the Carrington longitudes simultaneously in all three zones. The other half of the sun was dominated by divergence among the patterns. The hemisphere of convergence coincided with the active longitudes where 75% of all the sunspots, flares and coronal transients occurred (Hildner, 1977), while the other hemisphere was populated with large coronal holes (McIntosh et al, 1976; Zirker, 1976). Sunspot groups that occurred in the vicinity of the coronal holes were smaller and shorter-lived than the average (Speich et al, 1978). Convergence enhances flux emergence, divergence inhibits it.

6.3 Large-Scale Magnetic Patterns 1964-1974

A continuous view of the large-scale magnetic field evolution for nearly all of Solar Cycle 20 was constructed from zonal strips of 128 consecutive H-alpha synoptic charts (Figure 23). The time period covers from sunspot minimum late in 1964 to two years before the next sunspot minimum. Each chart is repeated three times in a staggered sequence from left to right in order to provide a more continuous view of the long-lived patterns with significant drift with respect to the Carrington longitudes.

Large-scale polarity patterns were clearly delineated in H-alpha at high northern latitudes even at the very beginning of the solar cycle, implying that the patterns could not have formed from the scattered remnants of sunspot magnetic fields. These patterns evolved very gradually, without any abrupt changes, as the sunspot number increased rapidly to solar maximum at the beginning of 1968. Most of the large sunspot groups of 1965-1966 in the northern solar hemisphere formed on the southwest border of the single, persistent, positive-polarity pattern in the northern high latitudes. This giant regular structure must surely reflect a large-scale convective pattern since it precedes and determines the location of important sunspot groups.

Magnetic-Field Patterns Solar Cycle No. 20

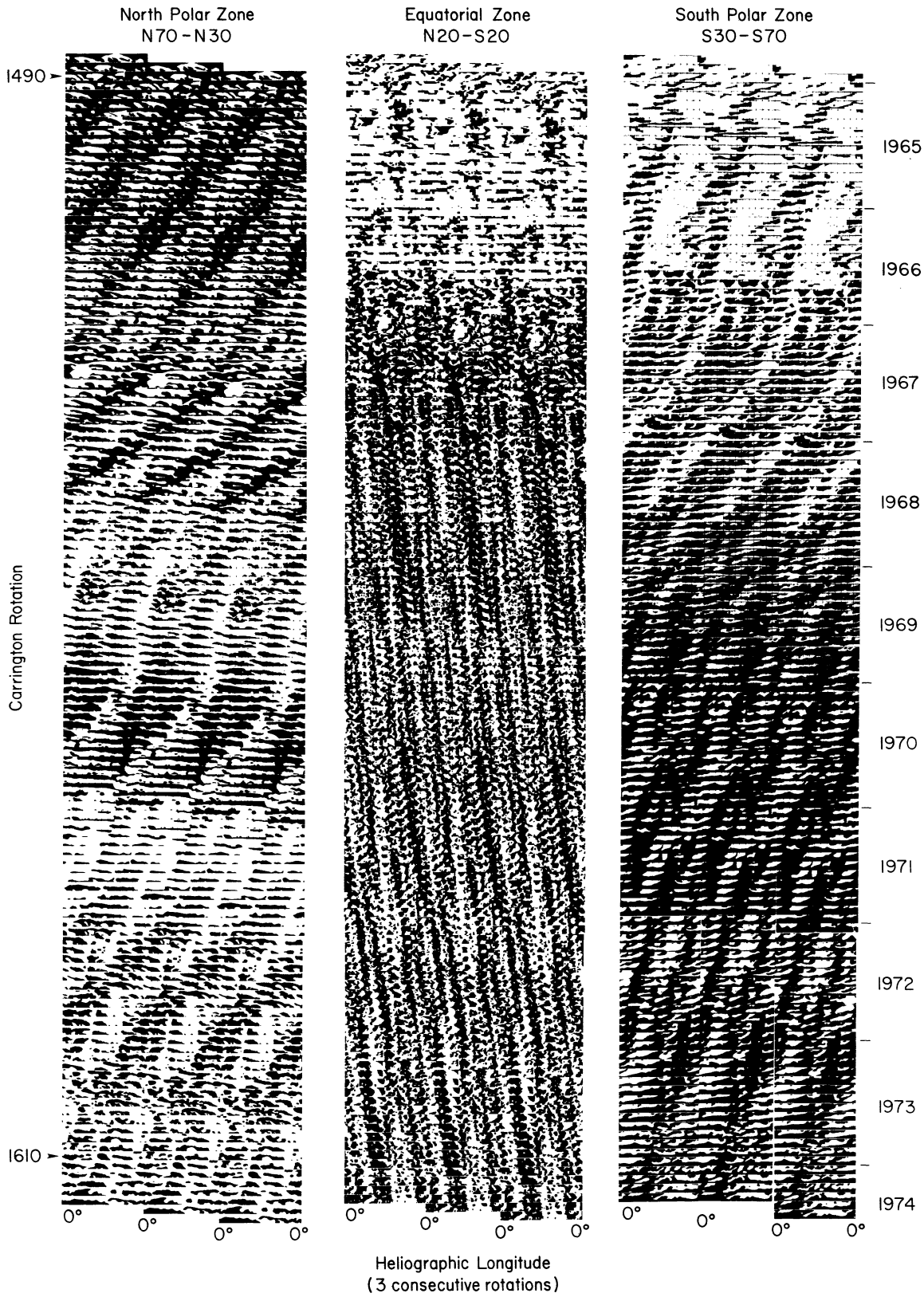


Figure 23

Large-scale magnetic patterns were virtually absent from both the equatorial and high southern latitudes until late in 1966 when the frequency of southern-hemisphere sunspot groups rapidly increased. It is normal for there to be no solar regions near the solar equator in the early years of a solar cycle, but the virtual absence of sunspots in the southern hemisphere two years after solar minimum was unusual. However, the formation of southern-hemisphere patterns, like those in the north, began before the appearance of large sunspot groups.

The first pattern to persist in the southern hemisphere was present, although ill-defined, in January, 1965. This pattern enlarged and became outlined with conspicuous filaments and filament channels following the formation of a small spot group in June, 1965 (Rotation 1495). The spot group never became large enough to develop penumbra, yet it was the largest and longest-lived spot group to form in the southern hemisphere since sunspot minimum a year earlier. Its strength relative to previous regions in that hemisphere was also indicated by the return of an extensive spotless plage on the following solar rotation. The western boundary to this pattern remained clearly defined from this time forward, becoming the eastern boundary to the polar crown gap (McIntosh, 1980) until the time of polarity reversal at high latitudes late in 1968. It is absurd to attribute the development of this large and long-lived magnetic feature to the occurrence of a single sunspot group of such small size. The emergence of this structure must, again, reflect large-scale convection beneath the solar surface, with the associated spot group a product of that emergence rather than being any part of the cause.

The black diagonal in the north polar zone and the white diagonal in the south polar zone during the first four years of the solar cycle map the polar crown gaps (PCGs) of the two hemispheres (McIntosh, 1980). These areas were of the same polarity as the solar poles in their respective hemispheres and formed an extension of that unipolar area to lower latitudes. Long-lived extensions of the polar coronal holes are centered in the PCGs. The diagonal patterns in Figure 23, representing the PCGs, are curved, showing a marked deceleration of rotation rate during the period of rise to solar maximum. The patterns in the equatorial zone show no curvature during the same interval, suggesting that the deceleration at higher latitudes resulted in a steady increase in large-scale shear from high to low latitudes. This changing magnitude of shear may be fundamental to the increasing frequency of sunspot formation.

The large equatorial cell associated with the August, 1972 and July, 1974 proton-flare sunspot groups can be seen in Figure 23 as one of three major dark structures that formed every two-and-one-half years. The other two also originated coincidentally in time and location with great sunspot groups. Additional examples and more detailed study of properties of the associated sunspots are needed before determining the necessary conditions for a sunspot group to signal the formation of a major new pattern in the large-scale magnetic fields.

6.4 Large-Scale Solar Circulation During SMY

An effective method of monitoring large-scale evolutions quantitatively is to perform subtractions of consecutive H-alpha synoptic charts. An application of this method to the study of active regions during the Solar Maximum Year(SMY) has demonstrated this to be a practical means of relating the large-scale dynamics to the localized dynamics within and surrounding sunspot groups. The general solar circulation over one-quarter of the solar surface, centered on one of the active regions studied in the Flare Buildup Study, is revealed in Figure 24. This display was generated by adding together the results of four pairs of synoptic chart subtractions. Vectors were drawn between the initial and next position of each neutral-line segment judged to be the same feature from one chart to the next. Only features with clearly-defined positions, and which were not disturbed by the effects of new active regions, were included. These features are most likely to reflect the true mass motions that are expected to cause the slow evolution of the large-scale patterns. Because of the large distances between large-scale neutral lines, vectors from the subtraction of just two consecutive charts give a very incomplete view of the large-scale flow. The addition of four consecutive subtractions, covering a period of five solar rotations, fills in most of the solar surface with vector information, although there is the risk that the nature of the motion field will evolve over this period of time.

The results of accumulating motions over five solar rotations showed remarkable persistence in the character of motions over this period of time. Both the direction and magnitude of vectors remained similar in any one particular area. There was confirmation, in general, of differential solar rotation, with vectors to the right equatorward of S20 latitude and vectors to the left poleward of that latitude. Of great importance is the large amount of variation in the magnitude of the vectors along any given line of equal latitude, and the variation in the line of zero motion with respect to the Carrington longitudes(rotation rate equal to the synodic period of 27.2753 days). If the values of the mean differential rotation applied to an instantaneous situation, the line of zero motion should be near, and parallel to, the line of S20 latitude. Instead, that line describes a sinusoidal wave with a significant displacement toward the solar equator at the longitude(180°) of one of the great sunspot groups of Rotation 1695. An area of motion with velocities appropriate to latitudes of S30 or higher had developed near latitude S15, while motions nearer the equator at this longitude were still near the norm for equatorial motions. This means that the position of the strong sunspots and flares of Rotation 1695 was an area of enhanced shear in the large-scale solar circulation. The fact that the sum of five solar rotations did not change this pattern indicates that the enhanced shear formed before the appearance of the sunspots, and may, therefore, indicate a causal mechanism for both the spot formation and the subsequent deformation of the region. The consequence of this shear was dramatically evident in the highly-sheared aspect of this area observed on Rotation 1696.

The line of zero motion was more convoluted in the position of the high flare activity, as if the mass motions were more turbulent there.

LARGE SCALE SOLAR CIRCULATION

Derived from
Neutral-Line Proper Motions

Carrington Rotations 1693-1697

Vectors Derived from Subtraction of Consecutive
H-alpha Synoptic Charts

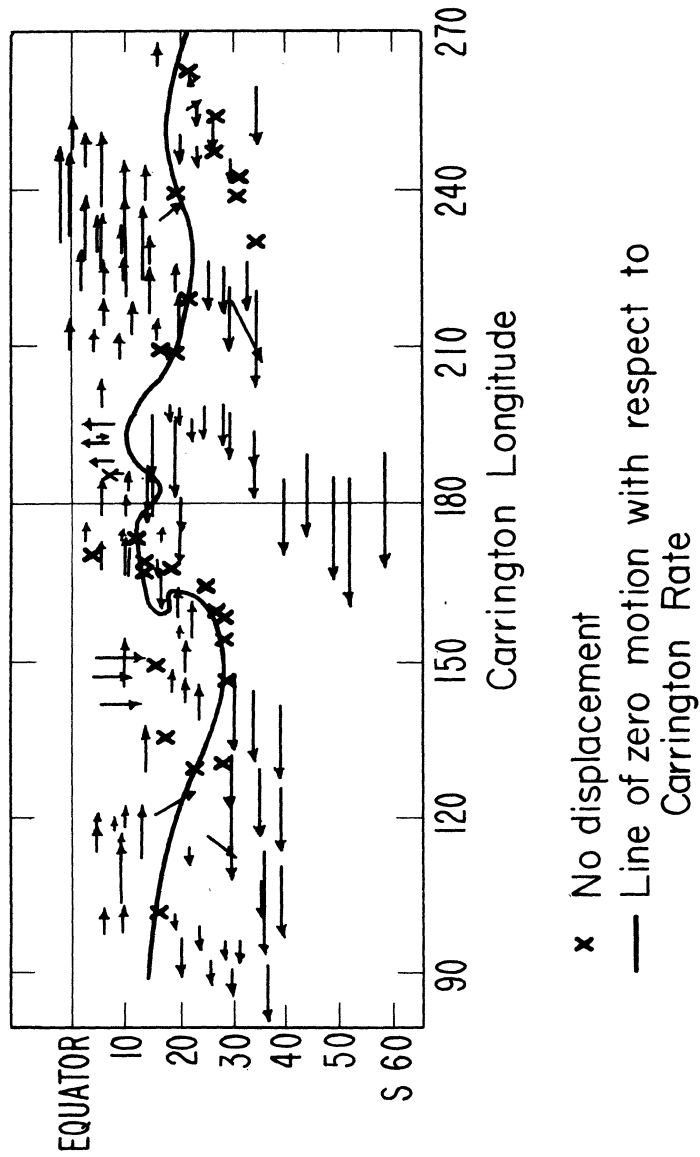


Figure 24

Vertical vectors mark the presence of meridional flow on either side of the position of the active region, with southward flow in the longitudes of greatest poleward displacement of the line of zero motion, and northward flow in the longitudes of greatest equatorward displacement. These meridional motions describe a large-scale, counterclockwise vortical motion centered on the location of the active region. This is the Coriolis effect expected from inflow, or from the observed differential rotation in the Sun's southern hemisphere.

6.5 Gleanings from the Atlas Annotations

The compilation of the atlas of H-alpha synoptic charts (McIntosh, 1979) required an exhaustive review of every feature on thousands of H-alpha filtergrams for the past 17 years. Recorded in the atlas annotations are many structural and evolutionary aspects of solar activity that are compelling evidence that large-scale processes generate and evolve sunspot groups. The evidence presented thus far in this review corroborates many of those notes, so it may suffice to simply list a few examples from them without illustration.

Large recurrent sunspots, such as the one discussed in Section 3.4, have rotation rates slower than the average for their latitudes. The mapping of such spots in context with the patterns of large-scale neutral lines reveals their involvement, in every case, with the large-scale patterns lying poleward of their positions. This association takes the form of a long filament extending from high latitude toward the spot and then partially encircling the spot with a pattern that suggests the Coriolis vorticity (clockwise in the South, counterclockwise in the North). The long-lived, large-scale pattern shifts eastward in time with respect to the Carrington longitudes, with the spot tending to follow along as if rooted to the giant feature.

Some sunspot groups have rapidly-separating leader and follower portions, with follower spots following after a high-latitude pattern and the leader spot remaining with an equatorial pattern. Filament disruption occurs when such divergence divides a large-scale magnetic pattern that is bounded by filaments.

Large sunspot groups sometimes occur in a series with each forming successively to the east of the previous group, defining a "rotation" rate for the source of magnetic flux that is like the large-scale magnetic pattern lying poleward of the series of active regions. The pattern moved in unison with the series of regions, and yet the pattern was clearly present for months before and after the sunspots.

The rate at which active regions disperse varies over a wide range that is not related to the size and intensity of the sunspots. There appear to be large-scale areas where regions disperse rapidly, perhaps coinciding with zones of large-scale divergence (Section 6.2). The location of long-lived spots has been tentatively identified as a large-scale area lying between zones of divergence and convergence.

The emergence of a new sunspot group near an older group often coincides with an increase in the rate of decay of the older group. The sunspots discussed in Section 4.1 grew simultaneously with the demise of a spot group a few degrees west of it. The shock wave emanating from the large flare of 28 August 1966 originated from a point between the two groups (Zirin and Lackner, 1969).

Large-scale neutral lines sometimes develop waves, or kinks, without the emergence of flux preceding this change. New active regions have been observed on the following rotation at the location of the deformation. These waves are probably the result of large-scale shear.

Mature sunspot groups that persist with their group axis inclined at a steep angle, or with the sign of inclination reversed from normal, appear to occur on large-scale neutral lines with orientations which also differ from normal; that is, oriented opposite that which results from differential rotation. This argues that large-scale patterns determine the sunspot group geometry.

There occur epochs of solar activity of a month or more when flux emergence is more prone to occur within existing sunspot groups, as if the large-scale process determining flux generation maintains a stable pattern for a longer-than-normal time.

7. Summary

This review presented evidence that the physics of sunspots is intimately tied to the dynamics of all scales of solar convection. The dynamics of giant-cell convection are manifest in the evolution of large-scale patterns mapped by H-alpha structures. Analysis of the large-scale velocity patterns surrounding a particular sunspot group can be achieved by subtraction of consecutive H-alpha synoptic charts. Such an analysis suggests the existence of zones of shear occurring in a sinusoidal wave encircling the sun, analogous to "jet streams" in a planetary atmosphere.

The formation of sunspots is related to local maxima in large-scale shear and convergence, with the details of location, size and geometry controlled by the supergranulation, mesogranulation and granulation scales of convection. Large-scale hydrodynamics plays a role in the evolution of sunspot group, as evidenced in the occurrence of vortical motions and patterns among the spots, and in the chromospheric fine structure and large-scale neutral lines surrounding the spots. The sense of vorticity follows a Coriolis law consistent with differential solar rotation and/or inflow.

The dispersion of sunspot magnetic fields does not create the large-scale magnetic patterns but may simply illuminate them, since it is now evident that the large-scale patterns predate the occurrence of large sunspots and are not greatly modified by them. The lifetime of sunspots is controlled by the large-scale evolutions surrounding the spots rather than by the smaller-scale convection. The dispersion of

sunspot magnetic fields appears to be a large-scale wave phenomenon rather than a random scattering.

8. Postscript

The observations used in this review and the manner with which they were treated reveal more than just the importance of dynamical aspects of solar activity. These observations have been available for decades, and the methods for their reduction are not dependent upon advanced technology. Their old-fashioned simplicity may have much to do with their being overlooked in recent years as vacuum towers, electronic imaging, satellite observatories and computerized data reduction have absorbed attention and resources. White-light and H-alpha observations should improve with application of these advanced technologies, but much of their value lies in their long record of solar activity contained in the archives of low-resolution observations. The steady accumulation of quality observations of solar activity over the entire solar disk is essential for research on the nature of the solar cycle, for study of secular changes in solar activity over decades and centuries and for understanding the rare, but important, varieties of solar activity. Such synoptic observing programs require long-term commitment of resources and management by patient and visionary scientists.

ACKNOWLEDGEMENTS

Praise and acknowledgement goes first to God for contributing opportunities for observations, for revealing knowledge before there was proof, for enabling perseverance during times of skeptics and poor support and for populating my life with encouraging friends. Much of this study would not have been possible without Frank Q. Orrall, who conceived of, and facilitated, the construction of the white-light patrol at Sacramento Peak. Significant support and encouragement have come from the late Robert B. Doeker, John W. Evans, Allen S. Krieger, Harold Leinbach, Edward C. Roelof and Peggy Shea. Donna Batchelor, Delores Nottage and Audrey Wilvert prepared this press-ready manuscript.

REFERENCES

- Ambroz, P., Bumba, V., Howard, R. and Sykora, J. 1971, in Solar Magnetic Fields(I.A.U. Symp. 43), 696.
- Baxter, W.M. 1960, J. Brit. Astron. Assoc. 70, 138.
- Beckers, J.M. and Schroter, E.H. 1969, Solar Phys. 10, 384.
- Born, R. 1974, Solar Phys. 38, 127
- Bray, R.J. and Loughhead, R.E. 1964, Sunspots, John Wiley and Sons Ins., New York.
- Bumba, V. 1965a, in Stellar and Solar Magnetic Fields(I.A.U. Symp. 22), 305.
- _____ 1965b, in Stellar and Solar Magnetic Fields(I.A.U. Symp. 22), 192.
- _____ 1976, in Basic Mechanisms of Solar Activity(I.A.U. Symp. 71), 47.
- Bumba, V., and Howard, R. 1965a, Ap. J. 141, 1492.
- _____ 1965b, Ap. J. 141, 1501.

- _____ 1969, *Solar Phys.* **7**, 28.
- Bumba, V., Howard R., Martres, M.J. and Soru-Isovic, I. 1968, in Structure and Development of Solar Active Regions [I.A.U. Symp. 35], 13.
- Bumba, V., and Obridko, V.N. 1969, *Solar Phys.* **6**, 104.
- Bumba, V., and Sýkora, J. 1972, Upper Atmosphere Geophysics Report UAG 24 Part II, NOAA World Data Center A, Boulder, Colo. 311.
- Carrington, R. C. 1863, Observations of Spots on the Sun, Williams and Norgate, London.
- Chevalier, R.P. St. 1919, *Ann. Observ. de Zo-Se* **11**, B-10-24.
- Chystiakov, V.F. 1962, *Soviet Astron.* **6**, 363.
- Cortie, A.L. 1901, *Ap. J.* **13**, 260.
- Dezső, L. 1966, Debrecen Heliophysical Observatory, private communication.
- Danielson, R.E. 1961, *Ap. J.* **134**, 275.
- Hildner, E. 1977, in Study of Travelling Interplanetary Phenomena 1977, M. A. Shea, D. F. Smart and S. T. Wu (eds), D. Reidel, Dordrecht-Holland, 3.
- Howard, R., and Edberg, S.J. 1973, *Solar Phys.* **28**, 73.
- Howard, R. and LaBonte, B. J. 1980, *Ap. J. Letters* **239**, L33.
- Kildahl, K.J.N. 1979, in Proceedings of Solar-Terrestrial Predictions Conference, (ed. R. Donnelly), C-166.
- Leighton, R.B., Noyes, R.W. and Simon, G.W. 1962, *Ap. J.* **135**, 474.
- Loughhead, R.E., and Bray, R.J. 1958, *Aust. J. Phys.* **11**, 177.
- _____ 1961, *Aust. J. Phys.* **14**, 347.
- Mattig, W. 1971, *Solar Phys.* **18**, 434.
- McIntosh, P.S. 1962, Geophys. Research Directorate Research Note AFCRL-62-890, Air Force Cambridge Research Labs, Bedford, Mass.
- _____ 1964, *Sky and Tel.*, **27**, 280.
- _____ 1969a, *Ann. I.Q.S.Y.*, **3** (ed. A.S. Stickland) The MIT Press, Cambridge, Mass., 40.
- _____ 1969b, Upper Atmosphere Geophysics Report UAG-5, NOAA World Data Center A, Boulder, Colo., 30.
- _____ 1970, Upper Atmosphere Geophysics Report UAG-8, Part I, NOAA World Data Center A, Boulder, Colo., 22.
- _____ 1972a, *AAS Photo Bull.* **1**, 12.
- _____ 1972b, in Solar Activity Observations and Predictions, (ed. P.S. McIntosh and M. Dryer), the MIT Press, Cambridge, Mass. 65.
- _____ 1972c, *Rev. Geophys. & Space Phys.* **10**, 837.
- _____ 1972d, *Solar-Geophysical Data, Descriptive Text*, NOAA World Data Center A, Boulder, Colo., 25.
- _____ 1972e, Upper Atmosphere Geophysics Report UAG-24 Part II, NOAA World Data Center A, 303.
- _____ 1973, Upper Atmosphere Geophysics Report UAG-28, Part I, NOAA World Data Center A, Boulder, Colo., 30.
- _____ 1977, *Bull. Amer. Astron. Soc.*, **9**, 330.
- _____ 1979, Upper Atmosphere Geophysics Report UAG-70, NOAA World Data Center A, Boulder, Colo.
- _____ 1980, in Solar and Interplanetary Dynamics [I.A.U. Symp. 91], 25.
- McIntosh, P. S., Krieger, A. S., Nolte, J. T. and Vaiana, G. 1976, *Solar Phys.* **49**, 57.
- Newcomb-Engelmann 1948, Populäre Astronomie, 8th Edition, Johann Ambrosius Barth, Leipzig, 276.

- Nolan, B. Smith S. and Ramsey, H. 1970, Solar Filtergrams, Lockheed Missles and Space Company.
- November, L.J., Toomre, J. and Gebbie, K.B. 1981, Ap. J. 245, L123.
- Schroter, E.H. 1962, Zs. f. Ap. 56, 183.
- Simon, G.W. and Leighton, R.B. 1964, Ap. J. 140, 1120.
- Simon, G.W. and Weiss, N.O. 1970, Solar Phys. 13, 85.
- Speich, D.M., Smith, J.B. Jr., Wilson, R.M. and McIntosh, P.S. 1978, NASA TM-78166, George C. Marshall Flight Center.
- Stickland, A.C., ed. 1969, The Proton Flare Project, (The July 1966 Event), Ann. IQSY 3, The M.I.T. Press, Cambridge, Mass.
- Svetska, Z. 1968, Solar Phys. 4, 18.
- Svetska, Z. and Simon, P. 1969, Solar Phys. 10, 3.
- Unsold, A. 1955, Physik der Sternatmosphären, 2nd Ed., Springer-Verlag, Berlin, 158.
- Waldmeier, M. 1947, Pub. Zurich Obs. 9, 1.
- _____ 1955, Ergebnisse und Probleme der Sonnenforschung, 2nd Ed., Leipzig, Geest. U. Portig.
- Warwick, C. S. 1965, Ap. J. 141, 500.
- Weart, S.R. 1970, Ap. J. 162, 987.
- Weart, S.R. and Zirin, H. 1969, P.A.S.P. 81, 270.
- Wilson, A. 1774, Phil. Trans. 64, 1.
- Wilson, P.R. and Cannon, C.J. 1968, Solar Phys. 4, 3.
- Wilson, P.R. and McIntosh, P.S. 1969, Solar Phys. 10, 370.
- Zirin, H. and Lackner, D.R. 1969, Solar Phys. 6, 86.
- Zirin, H. and Lazareff, B. 1975, Solar Phys. 41, 425.
- Zirker, J.(ed) 1977, Coronal Holes and High Speed Wind Streams, Colo. Assoc. Univ. Press, Boulder, Colo.
- Zwaan, C. 1965, Rech. Astron. Obs. Utrecht 17(4).
- _____ 1974, Solar Phys. 37, 99.