IMPLEMENTATION AND MAINTENANCE OF THE ALIGNMENT OF ACCELERATORS (PRESENTED AS ÌSINKING AND AGEINGÎ)

Michel Mayoud
CERN, Geneva, Switzerland

Abstract

The precise alignment of the components of particle accelerators, and of their experimental equipment, requires special techniques derived from geodesy. The increasing size of these machines and the demand for tighter and tighter tolerances have led to the development of special instruments and methods. The present needs, in terms of relative accuracy along the beam lines, are around 0.1 mm (rms)—and they are fully satisfied by this geodetic metrology. But nothing is absolutely stable and perfectly rigid, neither the ground nor the structures : various geomechanical forces and progressive changes of mechanical properties in some material affect the positioning of the components (hence the subtitle), and the alignment is to be maintained regularly. This paper describes the basic concepts and techniques used for that purpose.

1. INTRODUCTION

Nothing is ìabsolutelyî stable and ìperfectlyî rigid, everything can be moved and/or strained and distorted by forces. All human constructions are affected, even our planet itself is subjected to various parasitic movements, random constraints or cyclic deformations. Accelerators, like other objects inserted in the upper layers of the earthís crust, are moved up and down, expanded or shrunk and distorted while getting older.

At the last stage of their installation, the most critical components are positioned (ìalignedî) to within 0.1 mm (rms), and then the geometry is progressively altered and must be restored. Even not knowing why it is moving, one cannot keep ignoring by how much: according to their related effects on the circulating beams, all significant movements have to be detected and corrected. This metrology calls for special techniques, coming from geodesy.

Geodesy is an old applied science, devoted to the measurement of the earth and to "positioning" problems in general, and whose means are also used for the measurement of some large to huge objects—such as dams, bridges, industrial or scientific equipment, etc.—and such applications are often referred to as "micro-geodesy". As far as high precision is concerned, one can also speak of "geodetic metrology". The basic concepts and some of the special techniques used for the initial alignment of accelerators and for their maintenance surveys are presented in this paper.

2. CATALOGUE OF FORCES ACTING AGAINST MAN-MADE CONSTRUCTIONS AND MACHINERY

In addition to the general tectonics and its related catastrophes (earthquakes, volcanos, etc.), subsidence (ìsinkingî) and upheavals may be caused by geomechanical forces capable of moving and deforming constructions:

- local tectonics, oregeny (active faults and cracks);

- level of water tables, seas and lakes, position and size of glaciers - i.e. ìenvironmentalî load of the substratum;
- water content and pressure in the ground, locally or over the whole area (circumstantial or seasonal), in connection with the physical behaviour of materials with respect to the addition or the removal of water — i.e. hydrophily and swelling of clays, structural deformations according to pressure changes in some layers;
- compression/decompression effects due to changes in the ìlocalî load and pressures (uplifts and convergence);
- specific weakness of the substratum caused by superposing underground constructions (galleries, caverns, etc.);
- thermal expansion of the ground around a heated construction (thermal convergence);
- progressive change of mechanical properties (due to ageing) in construction materials.

3. BASICS ON POSITIONING, DIMENSIONAL AND PHYSICAL GEODESY

Astronomy and geodesy have been linked since the very beginning of rational observations of the universe. Cadastral surveying appeared very soon, when establishing boundaries and taxes on land property. Astro-geodesy, as a set of precise ìpositioningî techniques, developed with the needs of navigators and cartographers. Then ìdimensionalî geodesy was a major science during the dispute about the exact shape of the earth (1735-44). Soon after, the development of specific mathematics and the improvement of measurement techniques induced questions on the physics of the earth. ìPhysicalî geodesy opened a new field of investigation about gravity anomalies and their effects on measurements and data processing. Then came the era of ìspatialî geodesy, using satellites (or even quasars) as links across or over continents. The whole geodetic science is closely connected to geophysics and other earth sciences.

What does a geodesist's toolbox contain? For global geodesy, the panoply is :

- Angular measurements : theodolites (best accuracy = 0.1 mgon / 0.3 arc second);
- Distance measurements : Electronic Distance-Meters (a few 10^{-6} accuracy, maximum 10^{-7} when using two carriers and high frequency modulation);
- Levelling: optical automatic levels plus Invar rods (accuracy <1mm per $km^{1/2}$);
- Accurate positioning: differential GPS, i.e. Global Positioning System with phase measurements on the two carriers while tracking the satellites (nearly 10^{-7});
- Intercontinental links : Very Long Base Interferometry (correlating signals from quasars), Satellite Laser Ranging, refined GPS;
- Mathematical reference surface and volume and systems : IUGG (International Union of Geodesy and Geophysics) or WGS (World Geodetic System) ellipsoid(s), local datum(s) and local tri-dimensional systems;
- Physical reference surfaces and volumes and data : geoid(s), mass models, gravity field and spherical harmonics, free air models for obtaining orthometric levelling data;
- Related mathematics and appropriate software.

When dealing with geodetic metrology, most of the above is still necessary (except intercontinental links) but some very accurate instruments must be added :

- Invar wires and other length measurement devices (a few 0.01 mm accuracy);
- Special instruments for mono- or bi-axial off-set measurements with respect to a stretched wire or a laser beam (a few 0.01 mm accuracy);
- Alignment telescopes and targets, rules and micrometers;
- Inclinometers (10^{-6} rad) or horizontal pendulums (up to 10^{-9} rad);
- Hydrostatic levelling systems (up to a few µm resolution);
- Interferometric calibration baseline.

4. SURVEY AND ALIGNMENT TOLERANCES FOR ACCELERATORS

The specifications of accuracy, for alignment, are related to beam optics—i.e. to the magnetic elements of the lattice. Transverse errors in positioning are seen as imperfections of the guiding field: the particles no longer meet the theoretical magnetic field or gradient, and this creates a local perturbation of the motion—which is specially critical in focusing elements. Depending on the magnitude, location and distribution of these alignment errors, the resultant orbit may undergo deviations and oscillations of varying amplitude—with possible resonant effects in circular accelerators. Accordingly, tilt errors induce vertical distortions of the trajectory, and related tolerances must be also specified for this critical parameter.

The relative positioning of quadrupoles is therefore of major importance along particle beam lines, and this is the reason why the main criterion for precision is a relative and local one, leading to the best "smoothness" along the trajectory. To be correctly defined, this statistical criterion involves the consideration of a local trend curve, fitted to the actual data. The rigorous estimate is therefore the remaining dispersion of positions around this trend curve. Nevertheless, it can be roughly expressed as the standard deviation (rms) of the discrepancies on the radial or vertical position of each focusing magnet with respect to the adjacent ones, comparing actual sagittas to the theoretical (expected) values.

In former accelerators, with a rather large aperture, this criterion was much less critical and technically easier to manage. But with the progress of measuring techniques and the economical gains in reducing the aperture, it is commonly set to 0.1 mm—or even less for future linear colliders for which a few micrometers are sought in dynamic alignment systems.

But the *absolute* accuracy is not without importance: long-range errors in curvature (or straightness) may also induce oscillations of the beam and degrade the performances of the machine. For "small" accelerators—even with 200 m or 300-m diameter—it becomes confused with "relative" errors, affecting the local smoothness. One cannot neglect this correlation: there is no ìsimplifiedî metrology for small machines. Errors in geodetic control points will also induce deformations in the adjustment of the metrological reference network used for initial alignment. As a consequence, the best absolute geometry is also desirable for a good mastering of the whole alignment process.

When considering the technical means able to satisfy such requirements, it appears that the vertical control (height ordinates) is perfectly and easily ensured by appropriate levelling techniques. The radial control (plane co-ordinates) is much more difficult to obtain for such tight tolerances, and the complication is drastically increased by the huge size of some accelerators.

5. NETWORK STRUCTURES FOR RADIAL (HORIZONTAL) CONTROL

First of all, for each new accelerator project, the exact location is ensured by means of a surface geodetic network, which provides control points at appropriate places, via access galleries or pits to the underground infrastructures. The accuracy of this framework has, of course, a direct influence on the control of the absolute geometry of the machine to be built.

The design of the metrological control network has evolved with the size of projects and with the measuring tools used at different epochs. Various accurate instruments, commercially available, can be used for measuring angles (with theodolites) and distances (with electronic distance-meters). Measuring short or medium distances with the required accuracy can still be a problem, and the observation of misalignments around the accelerator is not a straightforward process.

Some special devices were developed at CERN for this metrology:

- the Distinvar, for length measurements from 0.4 to 55 m, using calibrated Invar wires, with an in-situ accuracy of $\sigma \approx 0.03$ mm;
- the wire offset device, for measuring the offset of a point with respect to a straight line provided by a stretched wire (Nylon, Kevlar or carbon fibre, up to 120 m), with an in situ accuracy of $\sigma \approx 0.03$ to 0.10 mm according to the wire length and observation conditions;
- the laser offset device where the reference line is a laser beam and having about the same accuracy, depending on environmental conditions.

When comparing the network structures designed in various HEP laboratories, they can be classified into three main categories:

1. Regular polygons with central point (or central kernel of points), well adapted to small or medium accelerators and ensuring a stiff control of the absolute geometry, for example the Proton Synchrotron (PS 200-m diameter) shown in Fig. 1;
2. Ring-shaped networks with a chain of large braced quadrilaterals, still ensuring a good control of the absolute geometry but without central point, as for the Intersecting Storage Rings (ISR 300-m diameter) shown in Fig. 2;
3. Ring-shaped networks with either a narrow chain of quadrilaterals or a simple polygonal contour, maintained by a few control points at access areas, as in the Super Proton Synchrotron (SPS 7-km circumference) or in the Large Electron-positron Collider (LEP 27-km circumference) shown in Figs. 3 and 4. In such cases, the span between control points is 1100 m (SPS) or 3330 m (LEP), and it becomes more and more difficult to master the flexibility of such arcs.

The PS network was measured with accurate distances (Invar) and angles, and the positioning of magnets was first ensured by the same means, producing non-homogeneous radial-error ellipses (from 0.2 to 0.6 mm rms). Using offset measurements all around the ring, instead of polar measurements, brought a significant improvement (0.2 ñ 0.3 mm rms) on magnets.

The ISR network was a pure trilateration network, well structured and accurately measured with the Distinvar, which produced very good results, despite the low redundancy of its design. The radial error, along a diameter, was only 0.4 mm rms.

However it is worthwhile concentrating on some critical aspects of large control networks. The size of projects may quickly reach a point where geodetic parameters must be rigorously included in the positioning data: the earth is an imperfect ellipsoidal volume, and a Cartesian geometry does not fit directly with spherical co-ordinate systems, altered by gravity anomalies. In addition, these flexible quasi-linear networks may have a stochastic "behaviour" which raises many problems at different stages of their initial measurement, and later on with their successive use for the metrology of the actual object to be aligned and maintained, i.e. the accelerator itself.

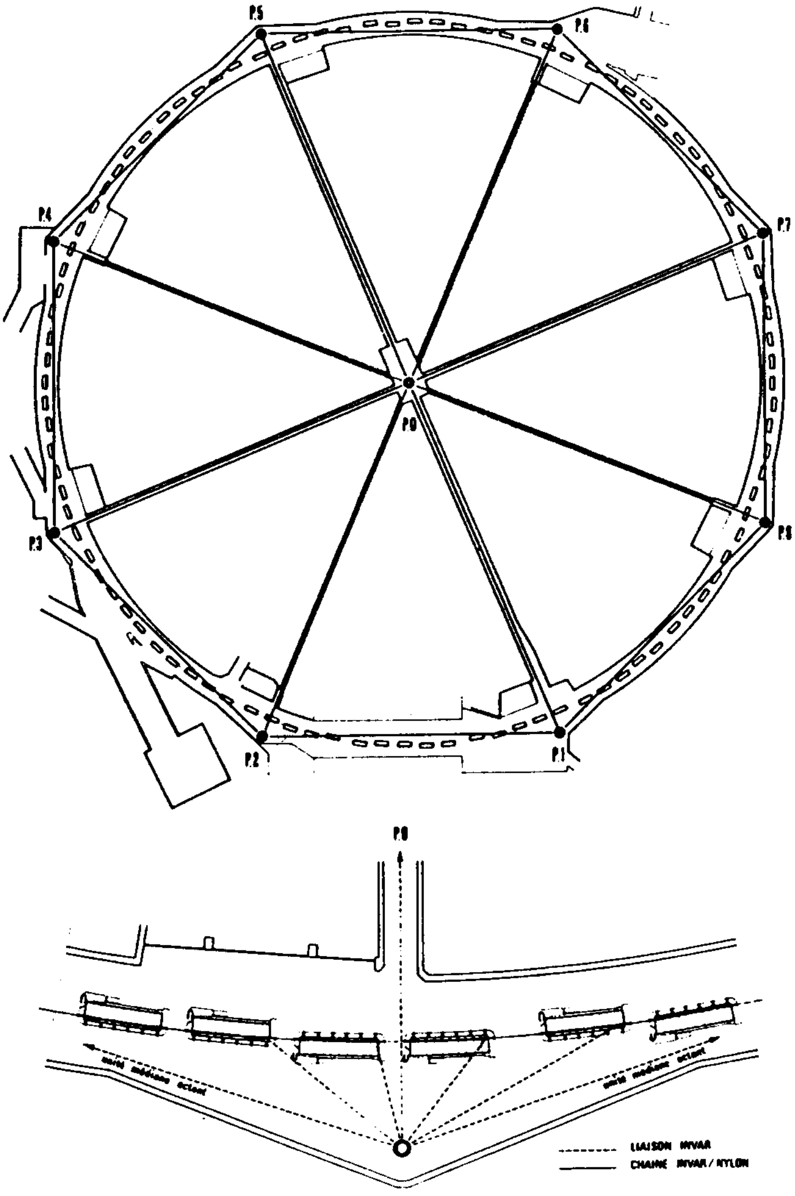


Fig. 1 Proton Synchrotron network

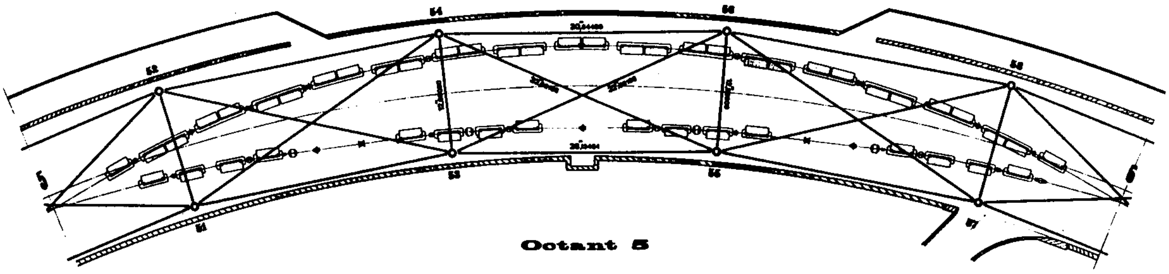
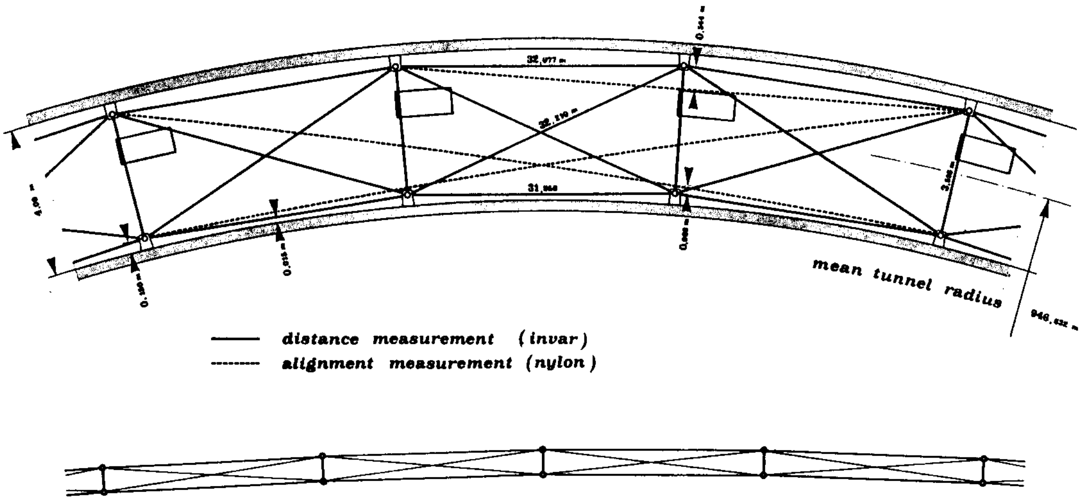


Fig. 2 Intersecting Storage Rings network



Exact scale of quadrilaterals

Fig. 3 Super Proton Synchrotron network

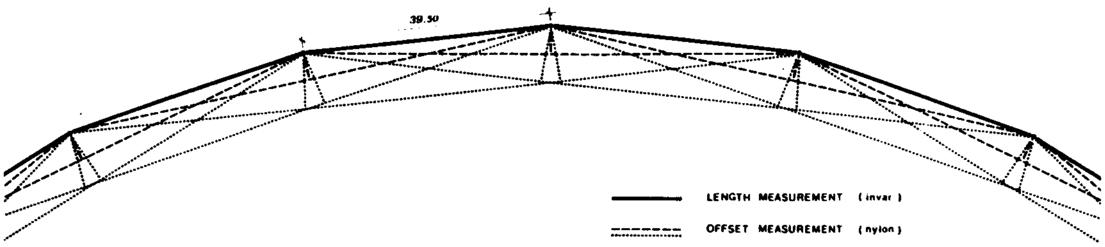


Fig. 4 Large Electron Positron Collider network

6. VERTICAL CONTROL

As mentioned, ìverticalî and ìhorizontalî (ìradialî) control is often processed differently and separately - except in 3-D triangulation/trilateration blocks, mainly used for the metrology of large experiments.

Vertical control is much simpler : it makes use of the well known technique of ìgeometricalî levelling, using optical levels and measuring height differences between successive points - thus forming ìtraversesî, ìloopsî when coming back to a starting point and finally a network when connecting loops. For reliability and quality, multiple readings plus forward and backward measurements provide redundancy, whilst loops (whenever possible) ensure a local check of the data. Care against systematic errors consists of making regular calibrations, ensuring symmetry (for cancelling refraction and earth curvature effects), looking at temperature effects on the instruments and on the measured structures, checking stability, etc. *Levelling is a rather simple process, complicated by many good reasons to do it wrongly!*

In modern instruments, there is no more ìhumanî optical reading: a CCD array and an ìencodedî staff (rod) do it. Along beam lines, in tunnels, the conditions of observation are very good and the rms accuracy of ìhigh precision levellingî can reach 0.3 mm per $\text{km}^{1/2}$ - with a rms error not greater than 0.04 mm in the height difference between adjacent quadrupoles (every 40 m in LEP). The closing error after the 27 km loop has often been (luckily) less than 1 mm!

For rather unstable sites (like ESRF), or for sensitive quadrupoles in low • sections (where misalignments and movements have amplified effects on the orbits), it can be interesting to set up a real-time levelling system for measuring the changes in vertical positions - combined with motorised jacks for correcting movements or resetting given positions. In the environmental conditions of accelerators, and for high precision requirements, the best technology is made of hydrostatic levelling systems (HLS) equipped with capacitive sensors (Fig. 5).

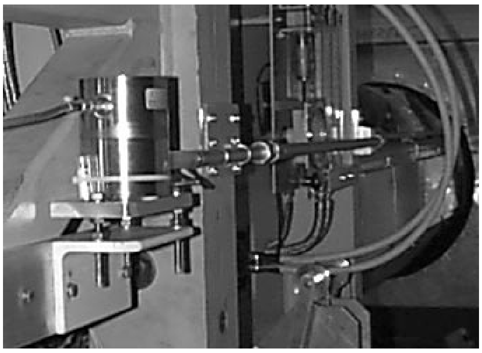


Fig. 5 HLS vessel on a low-β magnet

The principle of a HLS is very simple, but one has to avoid making either a thermometer or a barometer: only one input of air in the loop, and the smallest ìcolumnî of liquid. When temperature conditions are too much different between measurement areas, and when (in addition) the circuit of water has to snake up and down, the solution is to circulate the water through a tank sufficiently large to quickly obtain a homogeneous temperature before taking height measurements.

Other technologies are also possible for measuring the level of liquid:

- contact measurement, moving down a needle and measuring the displacement up to the ìtouchingî signal;
- ìparallacticî measurement, i.e. oblique beam of light (LED + optics) reflected on the surface and observed via a PSD;
- ultrasonic measurement (less accurate).

7. METHODOLOGICAL ASPECTS OF LARGE NETWORKS

The high-precision metrology of large networks demands not only that a maximum of accuracy and care be taken in the measurements but also that all relevant geodetic concepts are taken into account. This is mainly the case over large (or long) accelerators, of which tri-dimensional geometry is necessarily defined in a local geodetic system. For instance, at CERN, some major changes were introduced into the design of the SPS (2.2-km diameter) and LEP (8.6†km diameter) control networks for this reason.

First, in both cases, the computation of the theoretical XYZ co-ordinates of the machine has involved finer and finer consideration of the geometry of the earth. For the SPS, a spherical approximation was sufficient to express the effects of the earth's curvature in computing the Z ordinates, correcting the vertical "descent" of geodetic points along the shafts or properly tilting the magnets, in order to obtain a real plane in space. With the LEP project, which partly lies under the Jura Mountains, a further step has been to determine the vertical deflections generated by gravity disturbances, and then to express the separation between a reference equipotential surface and a reference (local) ellipsoid. This knowledge provides the necessary corrective factors to convert measured altitude into ellipsoidal heights in 3-D computations, to correct the co-ordinates of bottom points from the effects of vertical deflections or to reduce the gyro (azimuth) measurements for the difference between local horizon (physical plane) and geodetic horizon (local projection plane tangent to the reference ellipsoid).

One other change in the methodology is that repetitive measurements of the SPS or LEP control networks could no longer be thought of and managed as "absolute" surveys. For such long and flexible ring-shaped figures, the variations of the co-ordinates arising from different sets of comparable measurements may have no physical meaning. As mentioned, the trajectory of a beam within an accelerator is mainly sensitive to short-range errors. Long-range errors have less effect but are not negligible. In other terms, the figure must be smooth and this smoothing concept is fundamentally involved in a particular refinement process used for the first installation of a large machine and for any new partial or global survey when a re-alignment of components is to be carried out.

Finally it is worth mentioning that certitude in any accuracy problem cannot be acquired without a thorough knowledge of the stochastic behaviour of the measured networks. Although this statement sounds self-evident, it is in reality dependent on the method of estimating the actual errors and deformations which a network may undergo as a result of the random and systematic errors in the measurements, and also on the various constraints inherent to the chosen computational models.

8. STOCHASTIC ANALYSIS AND COMPARATIVE SURVEYS

Co-ordinates of geodetic networks are calculated by a least-squares adjustment of observations. The mathematical model (1, 2 or 3D) defines either a "free" network or a "constrained" one. The variances of adjusted parameters are derived from the Variance-Covariance matrix: Vx = s . N-1 and 1D error bars, 2D error ellipses or 3D error ellipsoids are derived from sub-matrices of Vx and their eigen vectors.

These estimates are not exhaustive, and the a posteriori variance of (groups of) observations may be altered: it depends on the redundancy and relative "strength" of each group, according to the network structure. As a complement, Monte-Carlo simulations allow an artificial generation of random and systematic errors, with controlled constraints. The effect of random (Gaussian) errors can be therefore correctly assessed, and the ìa posterioriî variances can be re-scaled. The deformations induced by systematic errors can be also isolated and identified. The whole gives true images of the distortions really suffered by a complex network. It provides "warning lights" to watch when actual measurement are made and processed.

For small accelerators, successive surveys must be processed in ìfreeî network adjustment and then superimposed (for comparisons) in a congruent transform : 2D or 3D Helmert transform with respect to theoretical co-ordinates. For large ones, the apparent flexibility of the arcs makes that only local comparisons are valid - after removal of non-significant differences.

Along quasi-linear traverses or networks, correlation is locally good but becomes very poor between remote points. The solution of the least-square adjustment produces therefore a rather ill-conditioned system. When repeating such measurements, the stochastic process of the random cumulating of normal (Gaussian) errors gives different profiles, with no physical meaning of the differences. Within the envelope of Gaussian errors, all ìwrongî lines have the same likelihood to be true—if no systematic errors alter the data, adding parasitic distortions.

The key problem is therefore to remove—analytically—these ìapparentî differences of stochastic nature - thus making the ìtrueî ones appearing really, as signals of significant movements. This must be made by ìtrend curveî analysis, completed with appropriate geometrical comparisons and relevant statistical tests (checking the signal/noise ratio). Applied to accelerators with the correction of the misalignments in order to reduce the dispersion, this process is referred to as ìsmoothingî.

9. RADIAL/VERTICAL SMOOTHING IN LARGE ACCELERATORS

When installing the machine components, the first determination of the control network gives the displacement vectors between their actual *rough* position and their theoretical one. As explained above, magnets are finally positioned around an unknown mean trend curve (one among an infinity) contained within the envelope of maximum errors. The polynomial degree of the curve depends on the redundancy, the overlap of measurements, and the bridge distance between control points.

The final relative errors are a quadratic combination of those of the network itself and those of the positioning, i.e. installation errors. Their statistical nature is essentially Gaussian: the aligned elements are randomly and normally distributed around this mean trend curve (Fig. 6).

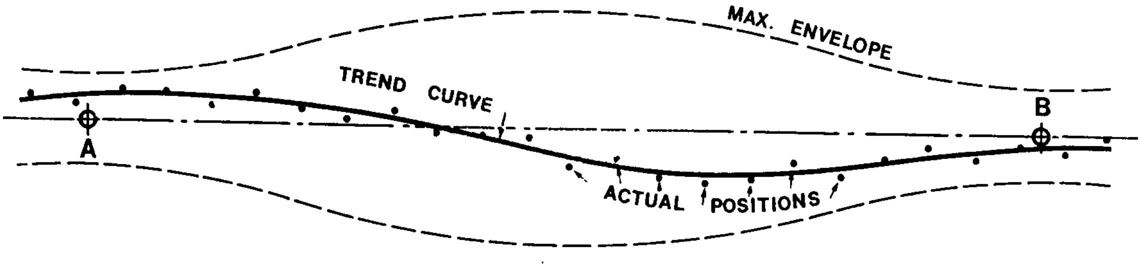


Fig. 6 Position of magnets with respect to theoretical orbit

As the major requirement for the geometry of an accelerator is that the relative errors must be very small ($\sigma \sim 0.1$ mm), a compulsory step is to check the installation by measuring and—if needed—improving the smoothness of the initial alignment. Then, during the exploitation, when making successive maintenance surveys of these long and flexible figures, absolute comparisons would be a nonsense and the differences between trend curves, corresponding to each survey, must be analytically eliminated. The state of the alignment is expressed by the statistical dispersion (rms) observed around the mean trend curve—after ìremovalî of the biases (large bumps and hollows) and using preferably the same algorithm in these successive comparisons. To restore the alignment to the required state, outliers have to be corrected and the global scattering has to be reduced up to the desired degree of smoothness.

The smoothing process consists of a set of radial or vertical measurements. For the radial case, it must be said that measurements cannot have the same span (and redundancy) as those of the metrological network, and the figure obtained—in co-ordinates—is less good, and more flexible. In any case, and for successive measurements as well, the problem is the difference between these distorted curves and the theoretical geometry. Each "image" of the ideal line has the same likelihood of being "true", within the envelope of errors, but is nevertheless different. This difference has (globally) no physical meaning, but local discrepancies or distortions may be the signal of a move, either for a single element or for a group, depending on the deformations of the supporting structure (floor and tunnel) due to geomechanical forces, and/or to some constraints along the machine (vacuum, dissipated energy, etc.).

Different concepts and methods have been tested for smoothing. The first one, used for checking the initial alignment, was a non-parametric method, which gave satisfactory results for this purpose but was not well adapted to the detection of movements. It even made us blind to some deformations! Parametric methods were then used with care.

As a matter of fact, polynomials fitted over such long lines and curves may induce correlations and constraints, which alter the image of the trend curve. Fourier analysis has also some drawbacks when choosing arbitrarily a given harmonic as the "best fit" and leaving cyclic errors around. Spline functions are rather heavy to handle and only piece-wise functions (attached arcs of polynomials, kept at a low degree) proved to be a realistic solution for a while.

Finally, a very satisfying method has been found. It consists in doing successive least-square fits of low-degree polynomials in a sliding window, shifted by steps all along the set of data. This concept allows one to retain the continuity of the trend curve—geometrically and statistically as well. It can be compared to a carpenter's plane used for smoothing an irregular plank: depending on the size of the tool and on the adjustment of its blade, one can obtain different qualities of planing with more or less waves on the wood.

In our algorithmic concept, abnormal offsets are located, and then removed if significant. This significance depends always on a signal/noise ratio at a given confidence level, and also to the choice of a given operational level. The adjustment of the blade is therefore a threshold, which defines a bandwidth beyond which data (i.e. actual offsets) are considered for a displacement back to the trend curve. The size of the sliding window can be fixed according to the desired degree of smoothing. Several iterations can be made, changing the parameters over the whole curve or in a given area when local discrepancies are observed at a higher degree of periodicity and have to be corrected more precisely. Such a method is very reliable, and flexible. Candidates for correcting movements are well identified, offset discrepancies are well quantified, and the whole work can be optimised according to the degree of "perfection" required.

CONCLUSIONS

It has been shown that nothing is stable and rigid, hence that everything is somehow floating—but not dramatically sinking—on the upper layers of the earthís crust, accelerators included. Nobody can prevent the evil but it must be now clear that everybody can cure it. Regular tests of geodetic metrology are necessary for the diagnosis, smoothing realignments make the potion for recovery, and particle orbits constitute the ultimate check 0n the good health of the geometry.

BIBLIOGRAPHY

W. Coosemans, M. Mayoud, E. Menant, D. Roux, Processing of the CERN Ring-shaped Geodetic Networks, CERN-CAST Workshop, 1-4 March 1977, CERN, Geneva, Switzerland

J. Gervaise, Applied Geodesy for CERN Accelerators (Part I), M. Mayoud, Computing and Analysis Methods (Part II), Seminar on High Precision Geodetic Measurements, Facolta di Ingeneria di Bologna, 16-17 October 1984, Bologna, Italy

CERN Accelerator School on Applied Geodesy for Particle Accelerators, 14-18 April 1986, CERN, 87-01

First International Workshop on Accelerator Alignment, 31 July - 2 August 1989, SLAC, Stanford, USA

Second International Workshop on Accelerator Alignment, 10-12 September 1990, DESY, Hamburg, Germany

W. Coosemans, MÈtrologie pour un Grand Collisionneur LinÈaire, XVII Colloque de l'AFT, June 1991, CERN, Geneva, Switzerland

J. Gervaise, M. Mayoud, Geodesy for Particle Accelerators, Advances of Accelerator Physics and Technologies - Chapter 16, Editor : H. Schopper, Advanced Series on Directions in High Energy Physics - Vol. 12, World Scientific

Third International Workshop on Accelerator Alignment, 28 September - 1 October 1993, CERN, Geneva, Switzerland