ALGORITHMS

E.Wildner
CERN

Abstract

The algorithms are the kernel of the problem to steer and shape particle beams. Depending on the specific conditions for a certain problem the data used by the algorithm has to be well prepared and the algorithm should be such that depending on the input information, numerical and other problems are avoided. The first part of the summary report, gives a simplified general view of the different problems and ideas that were presented. In the second part the different ideas related to what has been said in the first part are pointed out for each of the presentations.

1. GENARAL SUMMARY

1.1 Model based systems

Automatic beam control needs some kind of strategy. This strategy can be based on the existence of a model of the system or work in a way such that the desired control state can be achieved without predefined knowledge of the process. Models can be theoretical or empirical.

Control strategies based on a model obviously require that the system can be represented by a model to the detail needed by the control strategy or the required accuracy. Accelerators are designed from models, but it often appears that the real machine does not behave like the initial model predicts. If this is the case, investigations have to be made and the model revised or refined. Frequent problems that may be discovered are geometric problems (misaligned elements or measurement devices) and field problems.

Another important issue is the beam instrumentation, reliability must be good and accuracy must be within specifications. Some redundancy of information gives the advantage that the system can perform reasonably even if parts of the system fail (missing monitors for example).

The model based control mechanisms often lead to important improvements in the understanding and cleaning up of the process to be controlled.

When theoretical models cannot be set up or when the best theoretical model that can be found is not accurate enough, one can still make the control procedure automatic by using experimental transfer matrices. Either one can do this for the complete system or as a refinement of the initial model to control residual errors. Systems can also be linearized, assuming that the errors are small or that one can use an iterative control method.

A way to model highly non linear coupled systems is to use neural nets. Based on an experimental exploration of a sufficiently large domain of the control space, a neural net model and an adapted control strategy can be set up.

1.2 Systems where no model can be set up

If there is too little knowledge about the behavior of the system to be able to set up a model, one can use algorithms, which explore parameter space and try to find a good setting of the input parameters. This can be done in a random way or by using some clever strategy.

1.3 Data preparation

One of the critical issues in automated beam control is the quality of the data fed into the algorithms. During commissioning of the control algorithms, the measurement devices have to be checked and calibrated, the correct positioning in the beam line has to be verified for example. In spite of this preparative work, monitors may break down, noise may influence the readings and drifts may appear. Noisy measurement can be treated by harmonic analysis. Faulty monitors can lead to numerical problems, wrong corrections or unwanted corrections. Systematic effects on the orbit have to be removed, for example momentum errors.

Sometimes a "zero orbit" or trajectory does not correspond to the optimum working conditions. There are sometimes target orbits different from the ideal theoretical orbit which have to be identified and included in data preparation.

1.4 Correction strategies

Different strategies can be used in trying to find the correction. One might want to find dominant kicks; few effective correctors are wanted especially in the beginning of the correction procedure, since for example non linearities might be important if the orbit is far from the target. Global and local corrections have to be distinguished. Sometimes a limited part of the orbit needs to be corrected. Most of the algorithms can work on a limited region of the machine.

1.5 Matrix conditioning

The transfer matrix has to be conditioned to avoid numerical problems. Missing monitors, for example, can cause linearly dependent columns and the matrix has to be reconditioned. This is however time consuming and it is more interesting if one can anticipate the problem by not using correctors causing such problems. A very rigorous method to avoid problems with singular response matrixes is the Single Value Decomposition method (SVD). By this, singular values can be found and correctors corresponding to those can be removed from the available corrector set.

1.6 The algorithms

Generally, the algorithms minimize the distance, according to some norm, of the measured orbit to the wanted one. The MICADO algorithm is one of the most often used, especially for trajectories where it quickly gives a good result. It is based on a least squares minimization. MICADO first chooses the most efficient corrector for the minimization. By keeping this corrector an additional corrector is chosen that gives the best correction and so on. The disadvantage is that for a given number of correctors, the best subset of available correctors is not necessarily chosen. There are algorithms that help out of this dilemma, but one has to find tricks not to exceed wanted response times.

The SIMPLEX method uses the l_1 norm for minimization of the distance between the wanted and the measured orbit. SIMPLEX can include linear constraints or inequality constraints.

Complete least squares methods need non singular, well conditioned matrixes. It is important to check that that there is no overcompensation using too many correctors.

Genetic and evolutionary algorithms are well known techniques for numerical optimization. Interesting features of this technique are ease of development and independence of readout noise and set point drifts. Lack of speed can be compensated by a good choice of the set point generation mechanism and by choosing appropriate quality factors. A mixed type of modeling can be done where genetic algorithms are used for fine tuning of the optimization based on beam line models.

The basic idea of the genetic algorithms is to explore the parameter space to find an optimum using nature's way of producing new genetic sets. New correction sets are produced by using "selection", "reproduction", "crossover" and "mutation" on the "parent" corrector set to produce better performing "descendants" (better corrector sets). Offspring sets are produced not only from looking at the parent's genes but also from taking the mutation amplitude (or its efficiency, judged from the improvement of the correction) into account. A wide search for the optimum configuration of a given system is done by generating and selecting better and better sets of input parameters based on predefined quality factors.

1.7 Speed

The efficiency of the algorithms is very important since many of the correction programs work on line; examples will be given in the summaries of the talks..

1.8 Modularity and transportability

The methods described are well known and the task of beam steering and beam shaping is a very general minimization problem. This means that it is possible and profitable to write libraries and transportable and modularized software. The most time consuming part however, is the construction of the man machine interface. To reuse and share MMI software is inherently very difficult because of the choice of hardware and of special software products. Methods for data preparation can be shared, and also general methods for the matrix construction from machine layout information (data base description of the machine).

2. THE TALKS

2.1 Algorithms and Closed Orbit Correction package for SPS and LEP, W.Herr, J.Miles, CERN

The Closed Orbit Correction Utilities (COCU) program is an example of very general program for closed orbit correction. It is machine independent, has high performance and is extendable (highly modular). The program works from a sequence of commands that can be generated from the man machine interface. All knowledge of the machine comes from the data structure. A certain number of modules can be called that prepare the input data so as to ensure that the algorithms do not have numerical and other problems. There is a large number of different algorithms for optimization are available (MICADO, SIMPLEX, MINC2, PSINOM, SVD, Chebyshev...). Local and global correction criteria can be used. The package allows for very fast response, which is a crucial issue, especially in view of LHC where orbit corrections are needed with 10-50 Hz frequency.

For the LCH, modules that treat partially uncoupled machines have to be developed. Common elements for the two beams, ground motion compensation and long beam-beam interactions also have to be taken into account.

The COCU package has been used essentially for the CERN SPS and LEP.

2.2 Simultaneous matching of Dispersion and Twiss Parameters, M.Giovanozzi, A.Jansson, M.Martini, CERN

Dispersion control is important for large momentum spread beams. A linearized matrix is used together with the MICADO algorithm, which proposes a few larger corrections to gain in accuracy. A combined matrix is constructed to have both dispersion and Twiss parameter matching in the same optimization procedure. Units and sizes have to be chosen to give a well conditioned problem. The method has been tested in the TT2 transfer line between the PS and the SPS. By creating random error seeds and measuring the correction effect. If iterative correction is used, higher order effects can be neglected. Convergence was guaranteed to be within the limitations of the measurement devices.

The system is implemented as an open system stressing modularity and reusability.

2.3 Beam Transport and Optimization Tools based on Evolutionary Strategies, L.Catani, INFN Sezione Roma 2

The talk described how beam steering and shaping can be done using Evolutionary Strategies at LISA, the 25MeV linear super conducting accelerator of INFN-LNF.

For beam steering, one application has been shown, where one can clearly see that the optimization procedure, in spite of a non monotonic control space, finds a path to follow to achieve better and better correction. Starting from bad transport conditions, it took 6 minutes to achieve an rms value of the positioning errors comparable to the readout noise of the BPMs. Instrumentation noise and parasitic effects is a source of perturbation for the process, but did not prevent the evolution towards an optimized configuration. The algorithm can be refined by finding better selection mechanisms for the off spring. Defining the quality factor that best fits the system is another issue. For example by weighting of the contribution of the to the rms value, the algorithm can for example start at the beginning of the beam line as an operator would do. Another example was presented where for a constant energy spread, the alignment of the beam and the beam size at the end of the beam line are optimized. A clear reduction of the beam size is obtained when the beam is better aligned.

The method cannot be proven better than model based systems and no guarantee of finding always the same result can be given. However, it gives a "good enough for the purpose" result. The long execution time is a drawback that can be reduced only by faster computers and by optimizing the algorithm.

2.4 Tuning Knobs for the PS-SPS Transfer Line, G.Arduini, K.Hanke, CERN

Global betatron and dispersion matching in the transfer line between the CERN PS and the SPS has been performed based on a carefully checked theoretical model and measured input parameters. The Twiss parameters at the entry of the line can be determined from the measurement of the beam profile at three different locations in the line using SEM or OTR monitors. The dispersion and dispersion derivative is measured along the complete line and the first turn of the SPS and the initial values at the entry of the line are determined using a three-parameter fit. For the matched optics a tuning tool has been developed to correct for the unavoidable residual mismatch. It is based on the response matrix of eight quadrupoles in the matching section of the line which has to be inverted in order to compute the required correction. Since in general the matrix cannot be inverted, singular value decomposition (SVD) is used to recondition it. The resulting tuning tool has been experimentally tested during the 1998 run. Starting from a matched optics, the horizontal and vertical beta functions were detuned and the effect on the injected beam monitored by a mismatch monitor installed in the SPS. Both simulations and the measurements show that the tuning tool works fine for all beam parameters except the horizontal beta function for which a non-linear behavior was found. The studies will be continued during the 1999 SPS run.

2.5 Beam steering at ELETTRA, E.Karantzoulis, Sincrotrone Trieste

Beam orbit stability is a very critical issue in all third generation synchrotron light sources. Special care has to be taken in the case of ELETTRA, because of mismatched energy at injection. Global beam correction is done once per run or less. Local correction systems use selected BPMs near the quadrupoles in the straight sections and the appropriate correctors. Local correction is done every 5 minutes to keep the orbit within 5 microns rms and is totally automatic.

The COCU package has been used at ELETTRA and has evolved into Gloc which is more user friendly (COCU is "expert oriented"), and has an elaborated user interface. The algorithm used for orbit correction is based on SVD, which insures that corrector kicks give reasonable improvements. The algorithm has been used to correct the dispersion by combining steering and the dispersion matrixes. Creating bumps without perturbing the global orbit makes the local corrections. However some "bump leakage" might appear. New algorithms are being developed to avoid this to happen.

2.6 Beam Threading in the LHC, H.Grote, CERN

The multipole fields of the LHC dipoles kick off center particles. Simulations including the dipole field and alignment errors show that particles do not even make one turn without correcting the trajectory. This situation can be remedied by using a "threader" that calculates settings for two upstream correctors from the transverse position and the slope at the offending position, such that with this extra double kick, position and slope become zero by using a linear machine model and iterations. BPMs are installed behind each quadrupole and the correctors in front of them. The procedure was tested and refined for LEP. The results were that pickups with bad performance have to be eliminated, and that more than two monitors have to be used for each correction to avoid excessive excursion of the trajectory over a larger part of the machine. This means that one has to find the optimum trade-off between corrector bumps and orbit amplitudes. The intention is to implement these improvements (inclusion of the COFFEE program to eliminate bad pickups and a fit mechanism to bring the orbit down without creating large bumps).