THE ANALYSIS OF PHASE-SPACE TEXTURES IN ULTRA-RELATIVISTIC MULTIHADRON PRODUCTION

Peter A. Carruthers

Department of Physics
University of Arizona
Tucson, Arizona, 85721 USA

and

Institut für Theoretische Physik
Johann Wolfgang Goethe-Universität
Robert-Mayer Str. 8-10
60054 Frankfurt am Main
Germany

ABSTRACT

In this paper we discuss the progress made recently in the analysis of higher order correlations. Although we mostly consider multihadron distributions, the methods are of wide applicability. In addition, concepts related to fractals and wavelet analysis are discussed.

INTRODUCTION

It has become clear in recent years that an understanding of high order correlations among particles produced in nuclear collisions is crucial for the purpose of answering several fundamental questions in this interesting field of research¹.

There are at least two crucial aspects to the required analyses. First of all, the methods employed cannot depend on uncontrolled approximations. Rigorous statistical methods are essential^{2,3}. The second aspect to keep in mind is the continued

absence of theoretical methods that can give a good description of complex multi-particle dynamics, let alone the still more complicated hadronization process at energies for which production of hadronic secondaries dominates the statistical analysis of final states.

Because of this the attitude my own research group has increasingly focussed on defining a rigorous statistical analysis that is independent of modelistic bias. Our method is to work from the outside to the inside rather than the usual inside to outside approach in theoretical physics. In a sense, we develop the analogue of a phase shift analysis, to precisely characterize data which can then be studied for their dynamical content. It turns out that very interesting systematics are emerging from this approach. The next step is to understand their meaning.

The key problem⁴ now to be faced is simple. Can we really learn anything of fundamental significance from work in this field? A typical collision in the relativistic regime produces a variety of complex products, from "uninvolved" nucleons in the projectile and target, which jump into phase space, evading equilibration, as well as generating additional complications due to impact parameter effects. The difficulty is how to separate the components of the reaction products in a clean way that allows the proper separation of the physical components.

In lower energy collisions, proton/neutron liberation dominates, along with collective flows and quasiparticle excitation. At higher energies with particle production and secondary-dominated ultrarelativistic collisions, it is also essential to have unambiguous tools of analysis. In the latter case the dynamical situation is especially ambiguous. In the former case it has been difficult to incorporate correlations and fluctuations into the dynamical models.

The procedures described here are based mathematically on point distributions of a single species of particle. New experimental results which discriminate between like-sign and unlike-sign or all charge mixtures are extremely important and must be dealt with. Some results are already available, and have become more interesting in the light of new results on Hanbury-Brown/Twiss (Bose Einstein) correlations in various reactions.

Historically, the earliest efforts to describe large multiplicity events in cosmic rays led to statistical methods and the absorption of relativistic fluid mechanics. We call this approach the SHM (statistical-hydrodynamics model), summarizing a large body of interesting work by Fermi, Heisenberg and especially the work of the Russian (Landau) and Japanese "schools". This methodology exploits the fact that the lowest order description merely expresses energy- momentum conservation and an equation of state which allows LTE (local thermodynamic equilibrium) to close the equations of motion. If the reader is interested in the details of the history of the subject, I recommend my lectures at previous NASI meetings held at Peñíscola in 1990⁵ and at Ciocco⁶ in 1992.

An important feature absent from these early works is a proper description of fluctuations². These can be geometrical, stereotypically depicted by the concept of irregular, perhaps fractal, event-by-event hadronization geometries. Or they can (as commonly measured) be count probabilities or moments thereof of number distributions in various patches of phase space. These days we are subjected to many claims about "the" equation of state, without a careful discussion about those parts of phase space (including of course time and spatial coordinates) for which the equilibrium results should be relevant.

Although pessimism is premature, it is essential to think clearly about the many complications that must be addressed if we are to find decisive evidence for "the"

physics. In order to evade such complications various "signals" have been concentrated, typically emphasizing unusual properties of hot nuclear matter which should indicate the onset of, especially, the quark-gluon plasma or other new phases of hadronic matter. Here we can mention dilepton pairs, photons, strangeness and other flavor enhancements J/ψ "suppression", etc. (As far as I can see, no "new physics" is out in sight.)

Despite the suggestive nature of these alleged phenomena, there remains no unambiguous indication that something really new is going on in high energy nuclear collisions. What is missing, we believe, is a serious and systematic analysis of correlations and fluctuations in nuclear collisions. We also believe that many of these deficiencies have been remedied in recent work.

Closely related is the issue of quasiparticles in excited hadronic matter. Hydrodynamic schemes, or one-body style kinetic equations need extension in this direction to capture the physics.

WHAT IS THE BEST WAY TO DESCRIBE EVENT BY EVENT MULTIPARTICLE FLUCTUATIONS?

Honesty requires some candor about the experimental debris that can easily conceal important systematics of bulk matter that everybody hopes to see. Among these observables, many are explained by non-deep physics, laws of large numbers, statistical averaging, etc. The crucial point here is to segregate those phenomena which are trivial (after the fact!) from those suggestive of something more interesting. In typical collisions we expect the production of resonances, fragments, diffractive components (strongly dependent on collision geometry) not to mention the presence of many hadrons. The inter-relation of all these mechanisms requires sophisticated techniques of analysis.

The need for black-box Monte Carlo codes remains important for the "predictions" of and analysis of experiments we need to avoid the illusion that such methodologies are predictive theories. In addition "corrected" data sets inherit the intellectual errors of the codes. For this reason we emphasize here recent developments which are based on mathematical identities. Despite the existence of a substantial mathematical literature, the particular needs of current high energy and nuclear data analysis have required special ideas as the precision of the experiments has increased.

In order to formulate the problem clearly we consider collisions of hadrons, described as

$$h + h, \quad h + A, \quad A + A'$$

where h is an "elementary" hadron such as p , π , K , and A a nucleus.

We note the incomplete nature of the experimental analysis, without considering acceptance and other tricky detector issues. Assuming (wrongly) a unique center of mass collision energy, we know that for high energies there will be a substantial dispersion in the number of produced particles. In addition, most contemporary detectors are not sensitive to neutrals, a fact which has to be considered. In fact, even the plus and minus aspect is often ignored, being replaced by the question of whether the particle is charged or not. The influence of the cuts made in the data samples, made with respect to the above analysis, is scarcely trivial.

Immediately we can identify some important issues:

1) What are the best observables? Besides mere counts, the further labels such as momentum, flavor, etc. will be crucial. In particular, the behavior of collective flow phenomena depends on the correct choice of dynamical variables. This should be a productive area of research.

2) What information best describes the geometry of the typical collision? For example, it is natural to identify high multiplicity with small impact parameter, but this is only a gross measure. Further, it is reasonable to expect fluctuations in many variables, in order to consider the statistical variation of the phase space geometry of the emitting source.

3) What is the interplay involved in the construction of ensembles and their analysis? In particular, one needs to consider the strong effect of rare and concentrated "spike" events on the statistical accuracy of the moment analysis. These non-trivial problems have led us to modify bin-oriented data analyses to improve statistical accuracy for high resolution.

Now, we have sets of points. Take one example, with density

$$\hat{\rho}(x, s) = \sum_{i=1}^N \delta(x - s_i) \quad (2)$$

where x locates the measurement and s_i range over locations labeled 1 to N .

Of course, one example is not enough, in general. The particular example² is only that, an example, and must be complemented by many others. We have repeatedly explained the methods necessary to formulate these problems.

1) Fundamental problems in probability theory. In contemporary experimental reality, there is no true statistical ensemble, even for $10^4 - 10^5$ events. The reason for this is that rare but large fluctuations corrupt the data analysis at high resolution, at least for the currently popular variables.

2) As mentioned before, we must find the most insightful physical variables. History has led to a natural emphasis on number counts and their correlations⁷. The further breakdown into flavors and other indicators given by distinct particles is well known to be important. Energy and transverse momentum spectra and correlations among these variables may become a dominant theme in the near future.

COUNTS AND THEIR ANALYSIS

First we recall the connection between moments and correlations: in a patch of phase space it is not really difficult to measure counts and their moments. This is in contrast to the difficulty of measuring correlation functions. As we have emphasized, averages of correlation functions identically yield count moments. The inverse process is more difficult.

In order to understand the nature of fluctuations it is crucial to subtract background terms. This leads from density to "cumulant" correlations. The prototype is given by the 2-particle example

$$\rho_2(x, y) \rightarrow \rho_2(x, y) - \rho_1(x)\rho_1(y) \quad (3)$$

(Corresponding to this is the number count:

$$\langle n(n-1) \rangle = \langle n(n-1) \rangle = \langle n \rangle^2 \quad (4)$$

(These formulae assume that the count distribution relates to one species of "particle".) Now we can note that this and higher order analogues provide a precise test of statistical independence, in the sense of Poissonian statistics.

The advance which caught the attention of experimentalists was made by Bialas and Peschanski in 1986. They partitioned phase space (then just rapidity) into bins and averaged over them in order to improve statistics. Although discussions among experts have occurred about the best way to assess the data we simply express the basic formula for the "bin-averaged factorial moment" F_p :

$$F_p = \frac{1}{M} \sum_{i=1}^M \frac{\langle n_i(n_i-1) \cdots (n_i-p+1) \rangle}{\langle n_i \rangle^p} \quad (5)$$

averaged over M bins of equal size, labeled n_i .

According to elementary arithmetic, Poissonian counts lead to $F_p = 1$.

The second idea proposed was that there could be a scaling behavior, originally conceived to be present in the rapidity variable y . The initial writings over-represented these speculations, as we have explained elsewhere.

Clearly if the correlation functions scale (whether for densities, or better, cumulants), the F_p moments, or their cumulant analogues K_p , will also scale.

The history of this subject is clouded by the premature zeal to detect scaling. We can see this by looking at a standard graph of F_2 .

As is clear², the form of $F_2(\delta y)$ is a simple manifestation of ρ_2 . This follows from the identity (here written for a single bin labeled i)

$$F_2^i \equiv \int_{\omega_i} dy_1 dy_2 \frac{\rho_2(y_1, y_2)}{\rho_i^2(\delta y)^2} \quad (6)$$

where ρ_i is the average single particle density in the chosen bin ω_i of size δy . The dependence of F_2 on the resolution δy is then determined by the experimental knowledge of ρ_2 . It turns out that although ρ_2 is not translation invariant, a simple exponential or Gaussian with a single correlation length gives as good a description of F_2 as more sophisticated parametrizations of the data. Fig. 1 indicates the behavior of F_2 as a function of resolution. The decrease for large bin size is a direct consequence of the cutoff due to the correlation length. The straight lines indicate the scaling fits to the high resolution domain. For the analysis of higher correlations and the cumulant decomposition, we refer to ref. 2.

There is no space here to discuss speculations about the structure of higher order correlations, and their connection with the negative binomial count distribution which is pervasive in the count statistics of multihadron production. (The validity of this approach in nucleus-nucleus collisions is not yet confirmed.) It seems true that the linked pair approximation for the hierarchy of cumulant correlations with appropriate negative binomial coefficients works quite well. Unfortunately there is no unique dynamical explanation at present.

Fig. 1. Best fits for F_2^y moments for exponential correlation $k_2 = \gamma \exp(-|y_1 - y_2|/\xi)$ for UA1 ($\gamma = 0.669$, $\xi = 3.24$) and UA5 ($\gamma = 0.656$, $\xi = 2.95$) (solid lines). Dashed lines show best fits using a power law correlation $k_2 = \gamma(y_1 - y_2)^{-\gamma}$.

PROBLEMS WITH THE BIN ANALYSIS AND IMPROVEMENTS THERETO

The big advance permitted by the BAFM (bin averaged factorial moment) approach was that high order correlations could be sampled, though not directly measured, because of precise new data on number fluctuations. However, even in 2nd order correlations one can detect procedural problems. The moment analysis is especially interesting in several respects. The dependence on resolution is perhaps the most important feature. This gives an indication of whether there is a scaling regime. Second, the trivial fact that Poissonian noise components are removed in the BAFM procedure needs little further discussion. However the usual approach forgets that it is necessary to subtract lower order background correlations, which obscures the meaning of the results. For this reason we reminded people of the necessity to use cumulant correlations². As a consequence, it became clear that contemporary data involving nuclei have almost no correlations beyond second order at present energies. The real issue is simple. Are there interesting fluctuations to be found in RHIC? Will a new phase appear at RHIC energies? Does multiple scattering destroy most of the correlations that contain information about dynamical fluctuations?

A more serious problem created by binning arises from rare but especially interesting "spike" events. The mathematical problem is obvious once pointed out. The counting procedure displays a major defect with regard to amplifying moment fluctuations for high resolution. Suppose that we improve the bin resolution by a factor of two, so a particular event with 12 points in a bin of size $2\delta y$. Further suppose that we have 5 and 7 points in the split bins of size δy . Consider the third moment F_3 . For the $2\delta y$ bin we have $12 \cdot 11 \cdot 10 = 1320$ for the un-normalized moment. Splitting the bins we have $5 \cdot 4 \cdot 3 + 7 \cdot 6 \cdot 5 = 270$. Despite normalization issues, it is clear that

the bin moments are sensitive to fluctuations due to procedure rather than nature. These difficulties can be removed⁸ by modifications of the moment analysis.

By redefining the moment definitions it is possible to reduce the noise at high resolution. In lowest (second) order this coincides with the well-known analysis of Grassberger-Hentschel-Procaccia¹¹. It must be admitted that devilish problems associated with normalizations remain. In higher orders there are details not to be brought up here.

For this meeting I would point to the problem of CPU time required to digest events of high multiplicity and geometric complexity. Thanks to the hard work of my collaborators, Eggers and Lipa^{9,10,12,13}, still another modification of the sampling of points needed for the correlation analysis, which they call the "star" integral, allows realistic data analysis techniques for relativistic nuclear collisions in which hundreds of produced particles will be typical. The required CPU time becomes manageable in this framework.

The papers in press need to be consulted for detailed computational procedures.

MORE REMARKS

I have not spoken much about fractals, multifractals, and so forth although these concepts are always on my mind. Many of these techniques are based on models without background noise. Therefore the applications of standard formulas can be dominated by noise in phenomena like nuclear collisions. After many meanders, I came to appreciate a different approach, the wavelet method which seems to have been properly formulated especially by Belgian and French scientists. The wavelet transform acts as a microscope, localizing and magnifying the texture of the histograms. A good introduction to wavelet theory can be found in Doubechies' ten lectures¹⁶.

I consider this to be the tool of choice for texture analysis, for data compression and the simplification of many body dynamics. In addition I predict that in a year everybody will recognize the possibility to improve and economize on detector design using such ideas. In a recent paper Hakioglu and I discussed¹⁴ power spectrum methods of data analysis. However, despite my pleasure in constructing that work, the inability of Fourier techniques to simultaneously localize in space and time is a symptom of its weakness.

The lecture by Martin Greiner¹⁵ at this meeting gives a nice introduction to the ways in which wavelet transforms could be useful in nuclear and particle physics.

ACKNOWLEDGEMENTS

In addition to my collaborators in my research group, I have had the great pleasure to work with European colleagues in many countries. The support of the Alexander von Humboldt Stiftung, the GSI Darmstadt, the Institute for Theoretical Physics in the J-W Goethe University of Frankfurt and the Deutsche Forschungsgemeinschaft, have been especially helpful. For this I thank Prof. Dr. Walter Greiner and his many talented associates. In the USA my activities have been supported by the Department of Energy Divisions of Nuclear and High Energy Physics.

REFERENCES

1. XXIV International Symposium on Multiparticle Dynamics. Aspen, Colorado. September 12-16, 1993. Many key references can be traced from the proceedings, eds. M. Block and A. White, (World Scientific, Singapore, to be published)
2. P. Carruthers, H. C. Eggers and I. Sarcevic, Physics Letters B254 258 (1991). From this one can find references to earlier work by the Tucson group.
3. P. Carruthers, monograph on "Correlations and Fluctuations in Complex Systems", to be published.
4. L. Neise, "Quantum Effects in the Phase Space Description of Particle Production and Multiple Scattering", these proceedings.
5. P. Carruthers, "Scale Transformations, the Energy-Momentum Tensor and the Equation of State" p. 513 in the Proceedings of the NATO Advanced Study Institute "The Nuclear Equation of State", ed. W. Greiner and H. Stocker (Plenum Press, NY, 1990)
6. P. Carruthers, "The Quark-Gluon Plasma: Search for the Holy Grail", p. 669 in the NATO Advanced Study Institute "Particle Production in Highly Excited Matter", Il Ciocco, Italy, eds. H. H. Gutbrod and J. Rafelski, Plenum Press, New York (1993)
7. P. Carruthers and C. C. Shih, Intl. Journal Modern Physics A, 1447 (1987)
8. P. Lipa, P. Carruthers, H. C. Eggers and B. Buschbeck, Physics Letters 285B 300 (1992)
9. H. C. Eggers, P. Lipa, P. Carruthers and B. Buschbeck, Phys. Rev. D (in press)
10. H. C. Eggers, P. Lipa, P. Carruthers and B. Buschbeck, Physics Letters 301B 298 (1993)
11. P. Grassberger and I. Procaccia, Phys. Rev. Letters 50, 346 (1983); H. G. E. Hentschel and I. Procaccia, Physica D8 435 (1983)
12. P. Lipa and H. C. Eggers, "Unbiased estimators for correlation measurements", preprint, September, 1993
13. H. C. Eggers and P. Lipa, "Star integrals and unbiased estimators", preprint, September 1993 (Cracow Workshop on Multiparticle Production, April 1993)
14. P. Carruthers and T. Hakioglu, Phys. Rev. D45 4046 (1992)
15. M. Greiner, these proceedings.
16. I. Daubechies, "Ten Lectures on Wavelets", Soc. for Industrial and Applied Mathematics (1992)