ENERGY STATISTICS IN A FILAMENT MODEL OF 3D VORTICITY

STEVE NUCHIA AND CHAD WESTPHAL

ABSTRACT. The static energy distribution associated with a 3 dimensional vortex method is simulated and classified as a first order degenerate U-statistic. A fitting process is developed to find the parameters of the equivalent distribution predicted by U-statistics theory. The corresponding thermodynamic quantities: entropy, temperature and heat capacity are discussed.

1. INTRODUCTION

A current model of a three dimensional turbulent fluid uses the following idea. Characterize a fluid by a collection of tubes of constant vorticity populating a volume. By doing this one can study the interactions of these so-called vortex filaments by numerical simulation. Furthermore this group of vortex filaments can be treated as a microcanonical ensemble thus enabling the computation of various thermodynamic quantities.

Here, our goal is to (1) numerically simulate a turbulent fluid by randomly generating vortex filaments, (2) to compute the distribution of energy associated with a population of vortex filaments, (3) to classify the energy of this vorticity model as a first order degenerate U-statistic, and (4) to use this distribution to discuss the computation of thermodynamic properties of the fluid.

In section 2 we describe two direct simulation algorithms. The results of our simulations are described in section 3. Surprisingly, it appears that radical changes in the parameters of the algorithm result in a pure rescaling of the energy distribution.

In section 4 we introduce relevant material from the theory of U-statistics and argue that our energy statistic is a first-order degenerate U-statistic. Based on the limiting distribution predicted by the theory, a fitting process is developed using Fourier Analysis in section 5. A surrogate distribution is found using this analysis and compared with the observed distribution.

In section 6 we discuss the thermodynamic interpretation of the energy distribution curve. Concluding remarks appear in section 7.

2. SIMULATION ALGORITHMS

An appropriate random generation of a large number of vortex systems will represent all possible realizations of a turbulent fluid with the same vorticity structure. Admittedly, knowing the appropriate random generation that will completely and

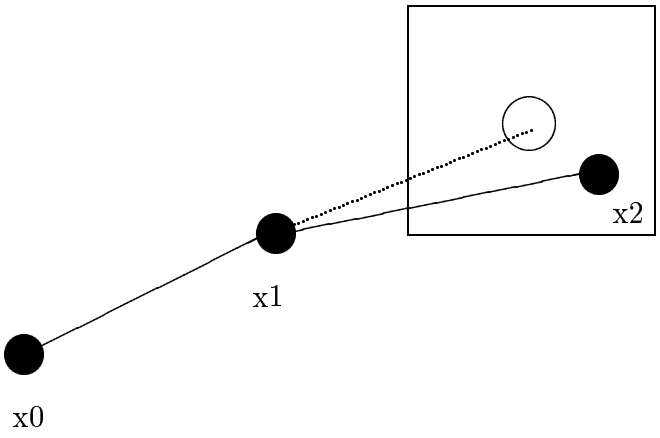

FIGURE 1. One step of the random walk by the first algorithm.

uniformly sample the space of the vortex filament arrangements is tricky. We avoid some of that complexity by studying the energy distribution of filament populations without attempting to tie them to a realistic physical system. Instead, we parameterized our simulation to allow for variation in such characteristics as population density, how kinked the filaments may be, length of the filaments, etc. Such parameters might be determined by the properties of a physical system.

The filament model is often used to simulate the time history of an evolving flow field. Here we are concerned with the distribution of filament representations in a static sense so we randomly generate populations of filaments and measure the energy associated with each snapshot. We developed two different algorithms, each motivated by the idea of building up a random sequence of vectors representing the segments of a filament. The invariance of the resulting energy distribution is independent of which algorithm is used.

The first algorithm produces a discrete parameterization of a vortex filament by generating a list of correlated points in $\mathbb{R}^3$ in the following way. Start at a uniformly random location chosen within a large cube centered at the origin. Choose another point one unit from the first. This comprises the first step. The next step is then a small perturbation (within a small uniform random cube) from what would be a unit step in the direction of the previous one. Figure 1 illustrates one step of this process.

The choice of the initial cube determines how densely packed a collection of filaments will be, and the choice of the perturbation on each step controls how curved a string will be. Step sizes are approximately one unit in magnitude – suitable values for the large cube are on the order of $50 \times 50 \times 50$ units – and the perturbation is typically on the order of unity in each direction. Also to ensure that the center of each filament is confined to the initial cube (and thus eliminating an unwanted bias in the calculation of energy) the algorithm begins at the starting location and works in both directions from the initial point.

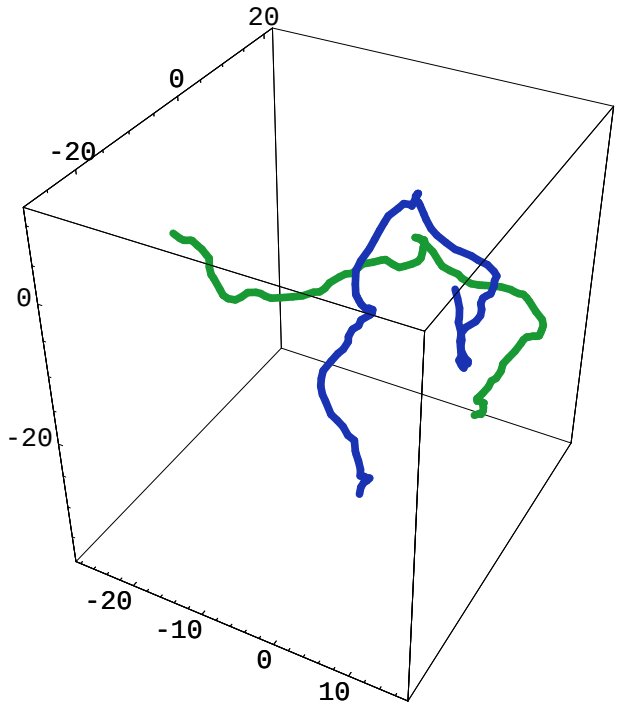

FIGURE 2. Two full-length filaments generated by the first algorithm.

This first algorithm is fairly simple and affords straightforward control of the parameters that define a turbulent fluid. However, computing the energy of interaction of a population of filaments using this representation is slow, in part because all computations are carried out in floating point. The second algorithm confines the endpoints of the segments to lattice points in a hexagonal close-packed (cannonball) lattice. In such a lattice the square of the distance between lattice points is an integer multiple of the fundamental bond length. In addition, each lattice point has twelve neighbors, each exactly one bond length away. Unlike the usual cubic lattice, there are bond direction pairs separated by an angle of 60° in a hexagonal lattice. This allows us to bend our filaments without introducing 90° intersegment angles. This idea was motivated by the problems addressed in [Cho93].

Details of this algorithm are available as a *Mathematica* notebook and an efficient C program. Briefly, a Markov chain generates a sequence of segment orientations chosen from the set of lattice bond orientations. Segments of length two (assuming unit bond length) are chained together to determine the relative midpoint locations of all segments. The filament is then translated so that the midpoint of the center segment is located at a point chosen randomly within some radius of the origin.

The algorithm is parameterized by the size of the starting sphere and the probability that the Markov chain will choose the previous direction at any step. The complement

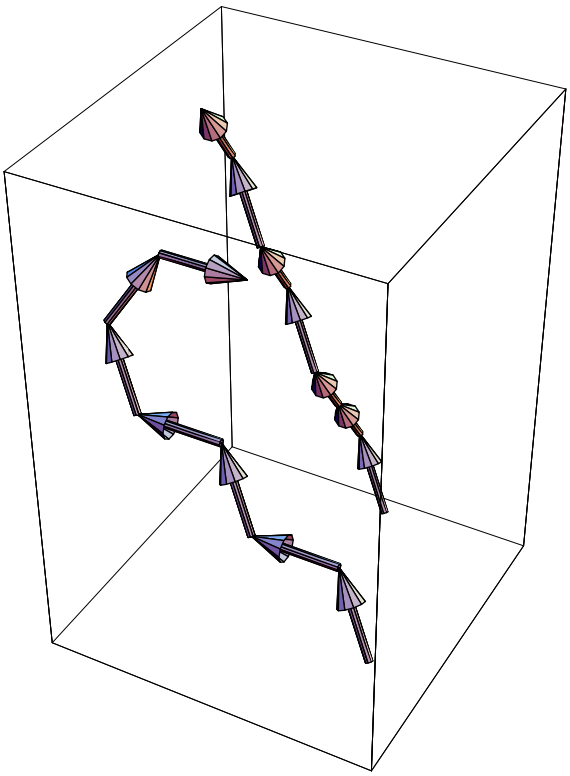

FIGURE 3. Two filaments in the hexagonal lattice with straightness parameter $P = \frac{1}{10}$.

of that probability is a measure of the curlyness of the filament. When a non-straight direction is to be chosen the algorithm picks one of the four orientations 60° away from the original direction with equal probability. Figure 3 is an example of two short filaments embedded together in space.

3. Simulation Results

Even though these algorithms are superficially quite different they produce energy distributions that are indistinguishable. In fact, varying the parameters of either algorithm produces a distribution of energies that is apparently just a rescaling of a common underlying distribution. In section 4 we interpret this result in terms of the theory of U-statistics.

Forty sets of 10,000 samples each were generated using the optimized version of the second algorithm with widely varying parameters. In addition, about 1500 samples were generated using the first algorithm and 10,000 with the *Mathematica* version of the second. Pairs of datasets were compared subjectively using the Quantile-Quantile plot method. The results are consistent with the hypothesis that their distributions are all related by a simple scaling operation.

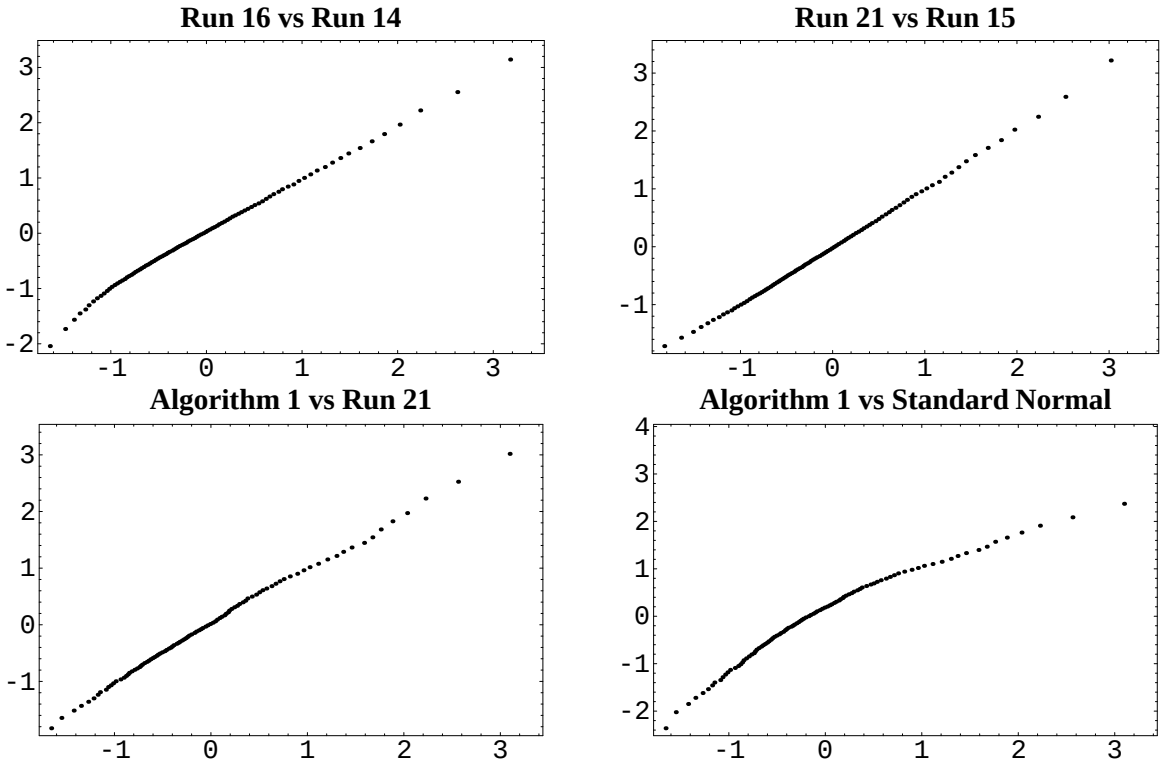

FIGURE 4. Consistency of Distributions. The fourth plot illustrates how the curve systematically deviates from a straight line when the distributions are not multiplicatively related.

In a Quantile-Quantile plot the coordinates of each point are given by the value of the ith quantile of the two distributions as i varies from zero to 100%. With sufficiently many data samples, the points will lie on a straight line if the two distributions are multiplicatively related. Here we have normalized the distributions by dividing each by its empirical standard deviation so that the slope of that line will be one. Figure 4 illustrates these results. A typical histogram is given in figure 5 and table 3 summarizes results of each run.

The forty runs represent a variety of parameter choices. Half of the runs were made with a population of ten filaments of 201 segment and the other half were made with twenty filaments of length 101. Straightness parameter values 60, 70, 85, and 95% were used along with central sphere radii of 25, 50, 100, 200 and 400 units. No systematic deviations from straight-line quantile-quantile plots were observed.

4. Energy as a U-statistic

The field of U-statistics is one that is both well studied and well understood. By classifying the energy of this vortex model as a U-statistic it is possible to capitalize

Run	#Filaments	%Straight	#Segments	Starting Radius	Q-Q Slope
A	20	n/a	100	50	1.00012
M	20	60	100	25	1.00626
1,	10	60	201	100	0.99896
2,	10	60	201	200	0.99036
3,	10	60	201	25	1.00343
4,	10	60	201	400	0.94946
5,	10	60	201	50	1.00835
6,	10	70	201	100	0.98409
7,	10	70	201	200	1.00767
8,	10	70	201	25	0.99466
9,	10	70	201	400	0.99130
10,	10	70	201	50	1.00046
11,	10	85	201	100	0.99595
12,	10	85	201	200	0.99538
13,	10	85	201	25	0.99365
14,	10	85	201	400	0.98158
15,	10	85	201	50	0.99963
16,	10	95	201	100	0.99988
17,	10	95	201	200	1.02144
18,	10	95	201	25	1.00558
19,	10	95	201	400	0.99627
20,	10	95	201	50	1.00291
21,	20	60	101	100	n/a
22,	20	60	101	200	0.99888
23,	20	60	101	25	1.00484
24,	20	60	101	400	1.01125
25,	20	60	101	50	0.99298
26,	20	70	101	100	1.00101
27,	20	70	101	200	1.00512
28,	20	70	101	25	1.00408
29,	20	70	101	400	1.00131
30,	20	70	101	50	1.01287
31,	20	85	101	100	1.01927
32,	20	85	101	200	1.01128
33,	20	85	101	25	1.00151
34,	20	85	101	400	1.00496
35,	20	85	101	50	1.01005
36,	20	95	101	100	1.00802
37,	20	95	101	200	1.01530
38,	20	95	101	25	1.00006
39,	20	95	101	400	1.01092
40,	20	95	101	50	0.99531

TABLE 1. Summary of the simulation parameters. Run A is the first algorithm, run M is the *Mathematica* version of the second. All others are from the C version. The last column give the slope of the linear least-squares fit to the Quantile-Quantile plot versus run 21. Run 21 versus standard normal has a slope of 1.06358.

on results from this theory. A U-statistic based on n measurements, U_n, is defined [Lee90] as the unique unbiased estimator of a regular statistical functional of degree k that has the form

$$U_n = \binom{n}{k}^{-1} \sum_{(n,k)} h(x_{i_1}, \cdots, x_{i_k}), \tag{1}$$

where h is the kernel of that functional. The sum is taken over all subsets of k elements from $\{1, 2, \cdots, n\}$ and $x_1, \cdots, x_n$ are random variables. Furthermore, if U_n is a U-statistic with zero mean and the expected value $\mathrm{Exp}[h(X, c)]$ is zero (with X random and c fixed) then the statistic is said to be first order degenrate and the normalized U-statistic, nU_n, converges to a distribution given by

$$nU_n \xrightarrow{\text{Dist}} \sum_{i=1}^{\infty} \lambda_i (Z_i^2 - 1), \tag{2}$$

with $\lambda_1, \lambda_2, \cdots$ being the eigenvalues of the integral equation [Lee90] (pages 79-80)

$$\int h(x_1, x_2) g(x_2) f(x_2) dx_2 = \lambda g(x_1), \tag{3}$$

and $Z_1, Z_2, \cdots$ being independent standard normal random variables. In equation 3 f is the PDF for individual filaments and h is the kernel of the energy functional.

To the extent that our distribution are in fact multiplicatively related, the linearity of equation 2 means that a single vector of "eigenvalue" ratios captures the invariant characteristics of the distribution. In terms of the eigenproblem, this means that changing the problem parameters so that the energy distribution is stretched in effect multiplies all the eigenvalues of the integral equation 3 by the same constant.

Each of the n strings of this vortex model are random and the energy of interaction is computed pair-wise. The energy kernel for two strings x_1 and x_2 (of lengths n and m) is given by

$$h(x_1, x_2) = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{s_{1i} \cdot s_{2j}}{\left| x_{1i}^m - x_{2j}^m \right|} \tag{4}$$

where s_{1i} is the ith segment of the string x_1 and x_{1i}^m is the midpoint of the ith segment of string x_1 [Cho93]. When a population of many strings is considered the energy of a single string with itself is negligible but when only a few strings populate a volume the self energy dominates. Since the former is the case of most interest the self energy is not considered in the energy kernel. Also it is found, using the simulation techniques described above, that for c a fixed string and X a random string, $\mathrm{Exp}[h(X, c)] = 0$. This is found to be statistically significant at the 0.05 (i.e., 95%) level. It is also interesting to note that the distribution of energies found holding c constant is symmetric while the distribution found using all random filaments is non-symmetric. Furthermore if we take the strength of each of the vortex filaments to be inversely proportional to the square root of n the self energies will

tend to a constant while the energy of interaction will then toward a value equal to half the value of the normalized U-statistic, nU_n, based on the energy kernel given in equation 4. Therefore, the energy of interaction given by the vortex filament model can be classified as a first order degenerate U-statistic of degree 2.

By admitting the expansion of equation 2, a first order degenerate U-statistic distribution can be reproduced if all of the eigenvalues are known. Direct solution of the eigenvalue problem is intractible, however, because the integrand is defined on a very complicated space - the space of all filaments in embedded in 3-space. Instead, we assume that a set of real eigenvalues exists and find an approximation of that set by a fitting process.

5. Curve Fitting

It is possible to construct, by trial-and-error, a set of λ values that generates a resonable approximation of the measured energy distribution. Fourier analysis permits a significant refinement of this method, but we have not discovered a well-formed optimization problem that will robustly estimate the required spectrum.

The PDF for one term of the limiting distribution (equation 2) with $\lambda_i = a > 0$ is

$$f_1(x) = \begin{cases} 0 & x < -1 \\ \frac{e^{-\frac{1+x}{2}}}{\sqrt{2\pi}\sqrt{1+x}} & x > -1. \end{cases} \tag{5}$$

This function is infinite at $x = -a$ and integrates to unity, so it can be seen as a distorted (smeared) Dirac Delta located at $x = -1$. The scaling properties of probability distribution functions require $f_a(x) = \frac{1}{|a|} f_1(\frac{x}{a})$ for all $a \neq 0$.

It is possible to compute the Fourier Transform of $f_a(x)$ symbolically. For $a = 1$ the transform is

$$F(\omega) = \frac{e^{-i\omega}}{2\sqrt{\pi}\sqrt{\frac{1}{2} - i\omega}}. \tag{6}$$

This formula is intuitively satisfying, as the numerator is the Fourier Transform of a Dirac Delta at $x = -1$ and the denominator corresponds to a mild low-pass effect. The scaling law for Fourier Transforms gives us the formula $F_a(\omega) = F(a\omega)$ for all $a \neq 0$.

From elementary probability theory we know that the PDF of the limiting distribution in equation 2 is the convolution of the PDFs of the contributing terms and therefore its Fourier Transform is the product of the transforms of the terms. We use this fact to estimate certain properties of the eigenvalue sequence.

Figure 6 shows a portion of the Discrete Fourier Transform of the empirical PDF (unsmoothed) associated with one simulation. In figure 7 we have zoomed in on the central portion of the transform. Data for higher freqencies is suspect due to statistical "noise"; to further refine the results would require a much larger set of data.

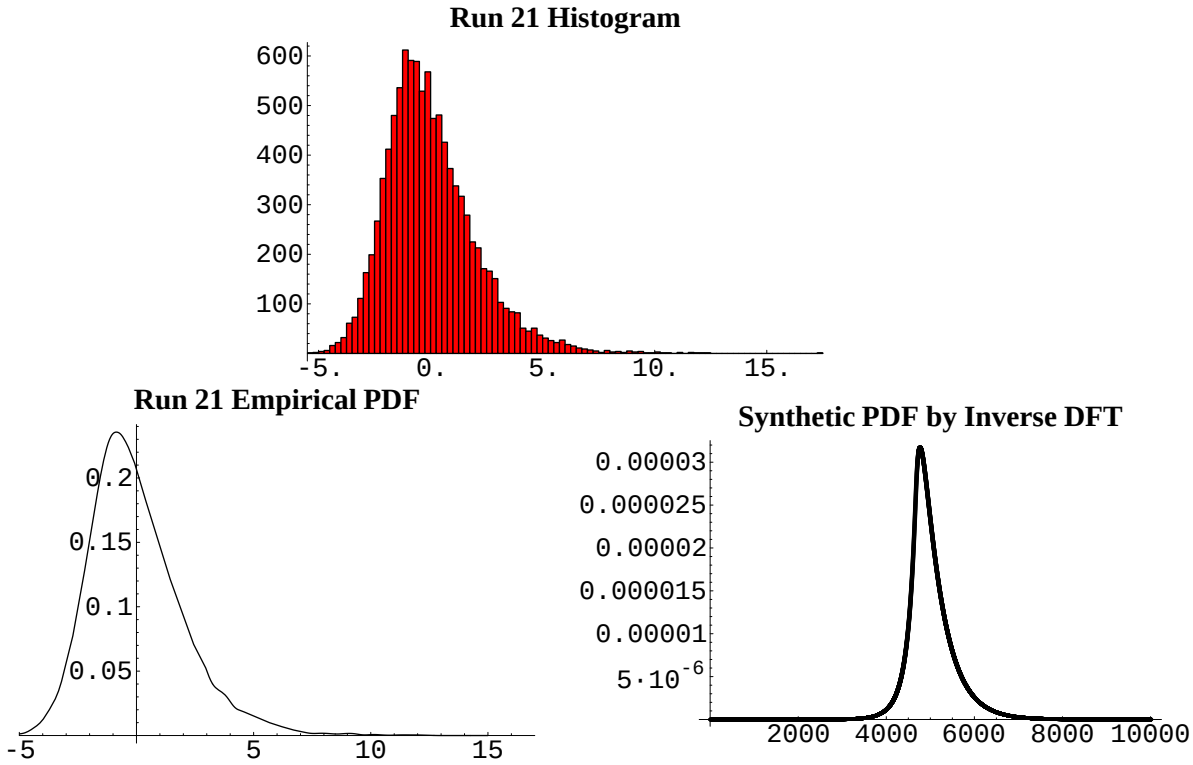

FIGURE 5. Energy histogram and empirical PDF for a typical run, together with a synthetic PDF obtained by Fourier Transform fitting. The empirical PDF is obtained from the raw sample set by digital differentiation and low-pass filtering. The synthetic PDF is the inverse DFT of samples from the transform generated using the eigenvalue set $\{\frac{11}{10}, \frac{9}{10}, \frac{-2}{3}, \frac{1}{3}, \frac{-2}{9}, \frac{1}{9}, \frac{-2}{27}, \frac{1}{27}, \frac{-2}{81}\}$.

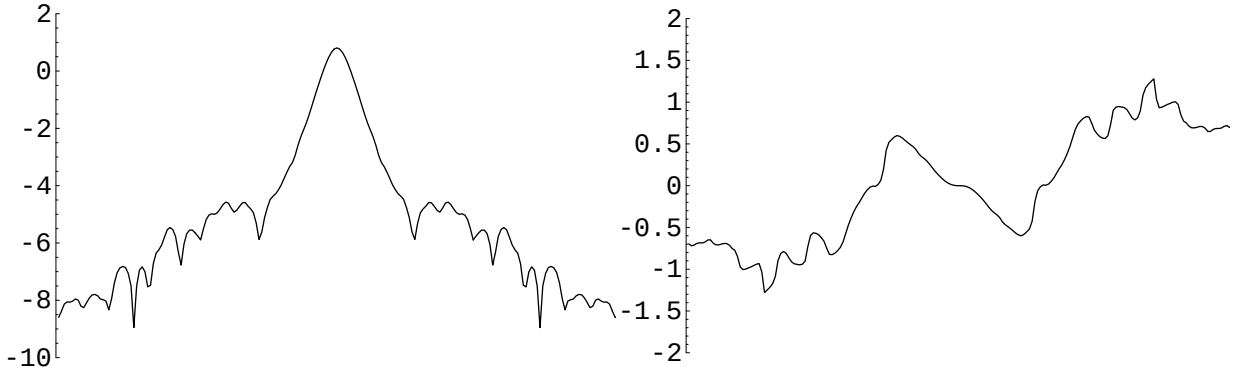

FIGURE 6. Log-magnitude and phase plot of the central portion of the Fourier Transform of the empirical PDF for run 21. Figure 7 is a magnification of this data.

Our fitting process concentrates on the structure of the phase curve. The phase (complex polar angle) of a product is the sum of the phases of the factors. The phase curve in our data exhibits several marked transition zones. We found experimentally that the location of the first zone corresponds to eigenvalues near one (the data have a distribution with most of its weight between -5 and 5). We therefore conclude that the largest eigenvalue is approximately one.

The phase of an individual component transform is given by

$$\arg F_a(\omega) = -a\omega + \frac{1}{2}\tan^{-1} 2a\omega. \tag{7}$$

Near $\omega = 0$ this is zero; far from the origin it converges rapidly to $-a\omega \pm \frac{\pi}{4}$. At frequencies that are low in relation to the scale of the largest eigenvalue the phase of the composite PDF is flat. At frequencies that are large in relation to all significant eigenvalues the phase is linear with slope proportional to the sum of the significant eigenvalues.

We see in the data transform plot that the phase curve is zero at frequencies below the first transition. Above the first transition there is a definite slope. The vertical offset between the two arms of the phase curve in this region is $+\frac{\pi}{4}$ so we know that there is one more positive eigenvalue than negative in the first cluster. The set of eigenvalues listed in the caption of figure 5 were chosen ad-hoc to satisfy these constraints. The idea is to pack an alternating geometric sequence into the transition zone so that both the sum and the net number of positive values meet the contraints imposed by the shape of the phase curve. The choice of a geometric series is arbitrary; we have no evidence that the eigenvalues of equation 3 actually have this structure.

At much higher freqencies, additional phase curve transition zones suggest the existence of a cluster of eigenvalues with a net negative contribution at least equal to the positive contribution of the first cluster. The curvature of the negative arm of the Quantile-Quantile plot in figure 8 corroborates this suspicion. We regard the data we collected as insuficient to resolve this feature with confidence.

The general shape of the PDF resulting from the fitted eigenvalue sequence can be seen in figure 5 to be similar to the empirical PDF. Also, a synthetic dataset generated using the imputed eigenvalue set in the limiting distribution is a reasonably good proxy for the original simulation: see figure 8.

6. Thermodynamic Interpretation

This filament model of a turbulent fluid is based upon the microscopic or small scale properties of a fluid, namely the vorticity structure. We would like to be able to relate these microscopic properties that can be simulated to macroscopic thermodynamic properties. This idea, the cornerstone of statistical mechanics, is one that we would now like to employ.

Consider an isolated system of a large number of "particles" occupying a fixed volume. Suppose that the energy of this system is known to be within E and $E + \delta E$.

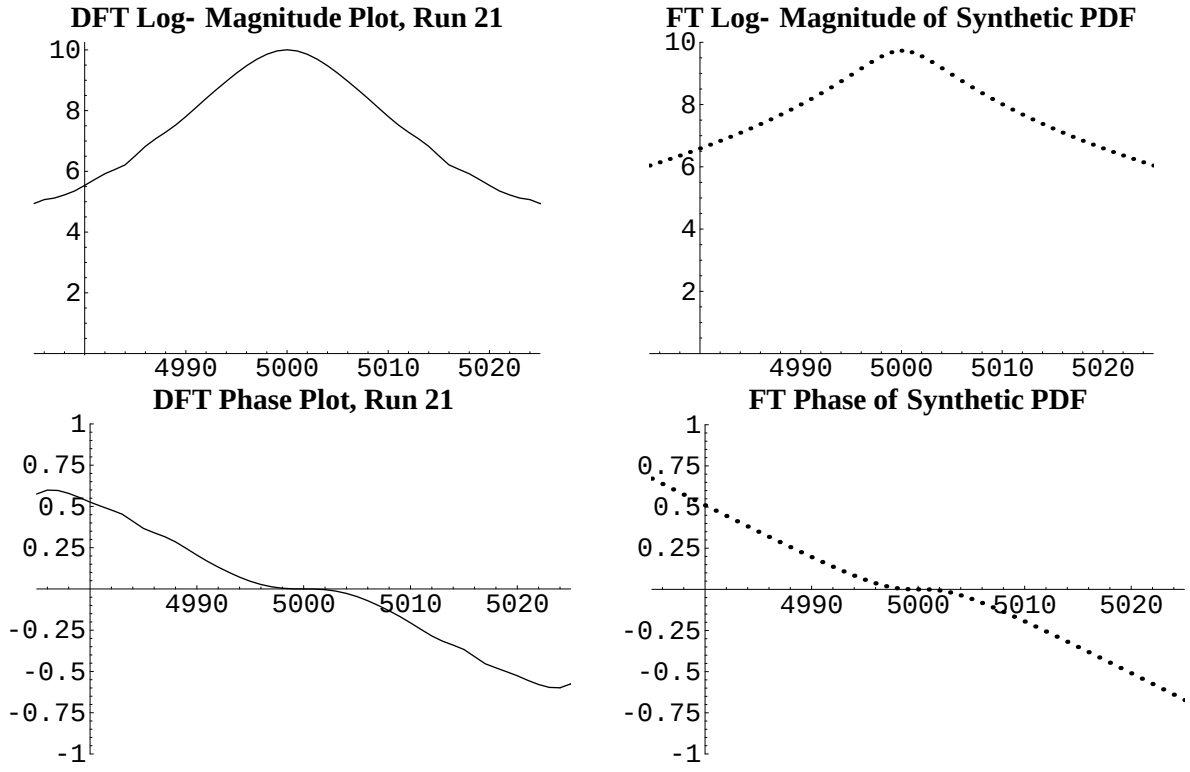

FIGURE 7. Fourier Transform of the empirical PDF of figure 5 near zero frequency, together with the corresponding region of the synthetic PDF transform. Phase scale is one unit equals 2π radians. See the caption of figure 5 for the vector of "eigenvalues" used.

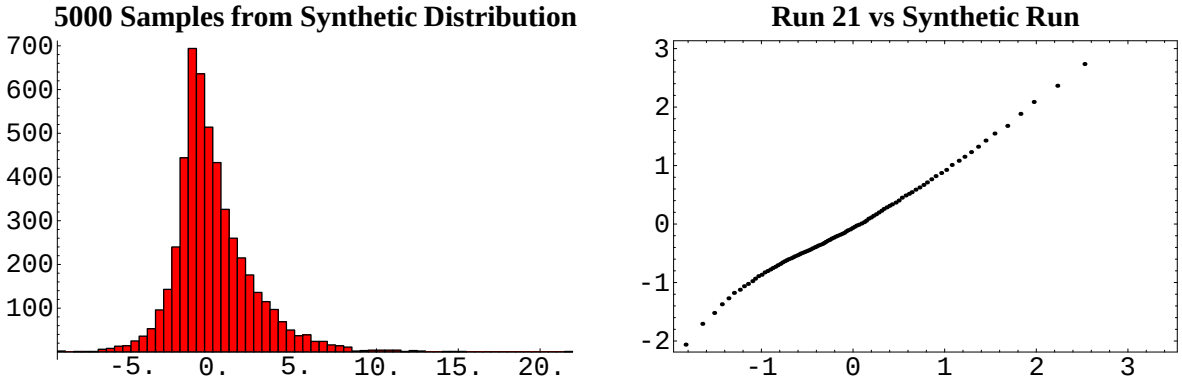

FIGURE 8. Histogram of data generated using equation 2 with eigenvalues from figure 5 and a Quantile-Quantile plot comparing that dataset with the data from simulation run 21.

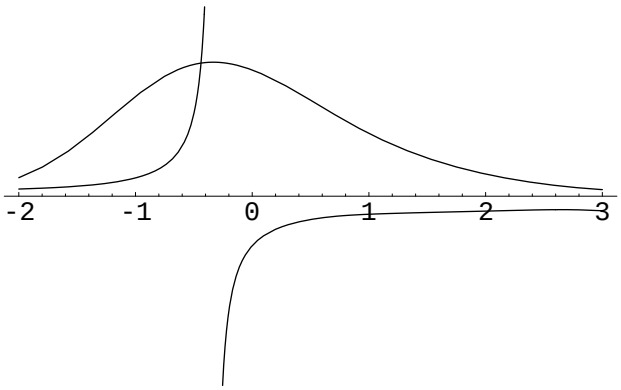

FIGURE 9. "Temperature" versus energy superimposed on the empirical PDF for run 22.

Studying the energy distribution of a large collection, or ensemble, of similar systems allows for a correlation to be made with the time average of the energy of a single given system. Such a hypothetical collection of systems is known as a microcanonical ensemble; see [Rei65]. Consequently, the energy and other macroscopic properties of a system of interest can be deduced by statistically analyzing the microscopic properties of a large number of systems. It turns out that many of the thermodynamic properties can be found from this distribution of energy which is often referred to as the partition function.

In the filament model simulations described above we take a fixed volume populated with a number of random filaments. We compute the energies of many of these systems and sort them into discrete bins. This histogram of energies is smoothed and normalized and taken as the probability density function for the energy of interaction of the filament model. Therefore let us take the energy distribution of the filament model found by simulation as $\Omega(E)$, the partition function. By definition the entropy of such a system is given by $S = k \ln \Omega$. Also known relations for temperature and heat capacity at constant volume, C_v, are given by $T = (\frac{\partial S}{\partial E})^{-1}$ and $C_v = T(\frac{\partial S}{\partial T})_v$. Describing these in terms of the partition function we get the following relations:

$$S = k \ln \Omega \tag{8}$$

$$T = \frac{1}{k} \frac{\Omega}{\Omega'} \tag{9}$$

$$C_v = k \left(1 - \frac{\Omega \Omega''}{(\Omega')^2} \right)^{-1}, \tag{10}$$

where k is the Boltzmann constant and derivatives are with respect to energy, E.

Now, these relations do characterize a turbulent fluid but some care must be taken when interpreting these properties. Since these were based on the energy of interaction of vortex filaments and not the kinetic energy of particles, the terms temperature and heat capacity do not hold the same physical meaning as they normally do. It is

common, however, to refer to the temperature of a vortex system. Figure 9, a plot of temperature vs. energy, shows the curious nature of these systems.

7. Conclusion

We simulated energy statistics of a vortex filament model. We were able to classify the interaction energy associated with this model as a first-order degenerate U-statistic. Using the theory of U-statistics a fitting process was developed in order to characterize the energy distributions by a set of eigenvalue ratios.

We found that even radical variation in the simulation parameters produced only a rescaling of the energy distribution. Also a connection was made between the static energy distribution of a large ensemble of states to the time average of the energy of a dynamic state. Corresponding thermodynamic properties were investigated.

Our data suggest the existence of additional structure in the eigenvalue set corresponding to the invariant distribution we found. A much larger set of simulation data would be required to determine that structure with confidence.

8. References

References

[Cho93] A. J. Chorin. Hairpin Removal in Vortex Interactions II. *J. Comp. Phys.*, 107:1–9, 1993.

[Cho94] A.J. Chorin. *Vorticity and Turbulence.* Springer, New York, 1994.

[CM93] A. J. Chorin and J.E. Marsden. *A Mathematical Introduction to Fluid Mechanics, 3rd ed.* Springer-Verlag, New York, 1993.

[CO91] L. J. Campbell and K. O'Neil. Statistics of Two-Dimensional Point Vortices and High-Energy Vortex States. *J. Stat. Phys.*, 65:495–528, 1991.

[Gre95] S.I. Green. *Fluid Vortices.* Kluwer Academic Publishers, Dordrecht, Netherlands, 1995.

[Lee90] A.J. Lee. *U-Statistics, Theory and Practice.* Marcel Dekker, 1990.

[Rei65] F. Reif. *Fundamentals of Statistical and Thermal Physics.* McGraw-Hill, New York, 1965.