C European Geosciences Union 2002

On the problem of optimal approximation of the four-wave kinetic integral

V. G. Polnikov^{1, *} and L. Farina¹

¹Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), Instituto Nacional de Pesquisas Espaciais (INPE), CPTEC/INPE, Cachoeira Paulista, SP, 12630-000, Brazil

*permanent address: The State Oceanographic Institute, Moscow, Russia

Received: 19 March 2002 - Revised: 2 July 2002 - Accepted: 25 July 2002

Abstract. The problem of optimization of analytical and numerical approximations of Hasselmann's nonlinear kinetic integral is discussed in general form. Considering the general expression for the kinetic integral, a principle to obtain the optimal approximation is formulated. From this consideration it follows that the most well-accepted approximations, such as Discrete Interaction Approximation (DIA) (Hasselmann et al., 1985), Reduced Integration Approximation (RIA) (Lin and Perry, 1999), and the Diffusion Approximation proposed recently in Zakharov and Pushkarev (1999) (ZPA), have the same roots. The only difference among them is, essentially, the choice of the 4-wave configuration for the interacting waves. To evaluate a quality of any approximation for the 2-D nonlinear energy transfer, a mathematical measure of relative error is constructed and the meaning of approximation efficiency is postulated. By the use of these features it is shown that DIA has better accuracy and efficiency than ZPA. Following to the general idea of optimal approximation and by using the measures introduced, more efficient and faster versions of DIA are proposed.

1 Introduction

It is well known (Komen et al., 1994) that a suitable description of a wind-generated ocean wave field is given by the two-dimensional wave energy spectrum distribution through the space and in time, $S(\sigma, \theta, \mathbf{x}, t)$. Here (σ, θ) is the angular frequency and the angle of propagation of the individual wave component, respectively; $\mathbf{x} = (x, y)$ is the geographical space coordinate vector and t is the time of evolution. In the theory, an evolution equation for waves is usually written in the form of a transport equation for the socalled wave action spectrum distribution in the wave vector \mathbf{k} -space, $N(\mathbf{k}, \mathbf{x}, t)$. In deep water and in absence of ambient

Correspondence to: V. G. Polnikov (polnikov@cptec.inpe.br or polnikov@mail.ru) currents it has the form

$$\frac{\partial N}{\partial t} + \left(\boldsymbol{v}_g(\boldsymbol{k}) \frac{\partial N}{\partial \boldsymbol{x}} \right)$$
$$= F\left(N, \boldsymbol{k}, \boldsymbol{U}, (\boldsymbol{x}, t) \right) \equiv IN + NL - DISS. \tag{1}$$

In Eq. (1), $v_g(k)$ is the group velocity of the wave component with wave vector k, F(...) is the source function describing the balance of energy for waves under consideration and U(x, t) is the local wind. Usually, the source function includes the energy-input term, IN, the quasi-conservative nonlinear wave-wave interaction term, NL, and the wave energy dissipation term, DISS (Komen et al., 1994). Correspondence between wave action spectrum, N(k), and energy spectrum, $S(\sigma, \theta)$, is given by the relationship

$$N(\mathbf{k})d\mathbf{k} = \frac{4\pi^2 g}{\sigma} S(\sigma, \theta) d\sigma d\theta.$$
⁽²⁾

In general, the wave vector \boldsymbol{k} is related to the frequency σ by the dispersion relation

$$\sigma^2 = gk \tan h[kD(\mathbf{x})] \tag{3}$$

where $D(\mathbf{x})$ is the local depth, and k is the module of the vector k. Later in this paper we shall restrict ourselves by consideration of the deep-water case when Eq. (3) reduces to $\sigma^2 = gk$.

In the following, the main attention will be placed on the *NL*-term. In several papers it was shown that the *NL*-term plays a principal role in ocean wave evolution (see, for example, Young and van Vledder, 1993; Komen et al., 1994). For this reason and due to a great mathematical difficulty involved in its study, research on this topic is ongoing for the last forty years, since the pioneering paper by Hasselmann (1962).

In his first paper on the subject, Hasselmann (1962) has shown that under some realistic assumptions, the rate of spectrum evolution due to four-wave nonlinear interactions is governed by the kinetic integral

$$NL = \frac{\partial N(\mathbf{k}_4)}{\partial t} \equiv T_N(\mathbf{k}) = 4\pi \int d\mathbf{k}_1 \int d\mathbf{k}_2 \int d\mathbf{k}_3 M^2$$

$$(\mathbf{k}_1, \, \mathbf{k}_2, \, \mathbf{k}_3, \, \mathbf{k}_4) \times \Big[N(\mathbf{k}_1) N(\mathbf{k}_2) \big(N(\mathbf{k}_3) + N(\mathbf{k}_4) \big)$$

$$-N(\mathbf{k}_3) N(\mathbf{k}_4) \big(N(\mathbf{k}_1) + N(\mathbf{k}_2) \big) \Big] \delta \big(\sigma(k_1) + \sigma(k_2)$$

$$-\sigma(k_3) - \sigma(k_4) \big) \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4)$$
(4)

Here $(k_i = 1, 2, 3, 4)$ are the wave vectors of interacting waves, $\sigma_1 = \sigma(k_i)$ are the corresponding angular frequencies of the wave due to dispersion relation, $T_N(k)$ is the nonlinear transfer of wave action, and M (...) are the matrix elements describing an intensity of interaction of four waves. The delta-functions in Eq. (4) assure that the four interacting waves should meet the following resonance conditions

$$k_1 + k_2 = k_3 + k_4, \tag{5}$$

$$\sigma_1 + \sigma_2 = \sigma_3 + \sigma_4. \tag{6}$$

A joint solution of Eqs. (5) and (6) defines a special resonance 3D surface in the 8-dimensional k-space. In a discrete representation, this surface give rises to a set of 4-wave configurations for wave vectors k_i contributing to the real non-linear transfer of wave energy among waves.

Due to symmetry properties of the matrix elements M, the nonlinear transfer formally conserves the total wave energy

$$E = \int N(\mathbf{k})\sigma d\mathbf{k},\tag{7}$$

total wave action

$$A = \int N(\boldsymbol{k}) d\boldsymbol{k},\tag{8}$$

and total wave moment

$$\boldsymbol{M} = \int N(\boldsymbol{k})\boldsymbol{k}d\boldsymbol{k}.$$
(9)

All these features of the kinetic integral result in specific properties of the real nonlinear transfer $T_N(\mathbf{k})$. The properties of the *NL*-term in a ocean wave model are dictated by the properties of the kinetic integral (4).

The problem of describing the *NL*-term can be divided into two aspects:

- 1. A theoretical study of the kinetic integral properties;
- 2. The implementation of the theoretical results into a practice of wind wave numerical modeling.

The explicit analytical expression of the integrand in Eq. (4) is rather complicated (see, for example, Hasselmann, 1962, or, in more convenient form, Crawford et al., 1980). In addition to this, the multifold integration in Eq. (4) is to be carried out on a specific 3D-surface in the 8-dimensional k-space with a singular locus. For these reasons a theoretical study of the kinetic integral properties is a very difficult task in the wind wave theory. This task gave rise to strong efforts

of numerous investigators, aimed to find the real properties of the nonlinear energy transfer among surface gravity waves (Zakharov and Filonenko, 1966; Webb, 1978; Masuda, 1980; Hasselmann et al., 1981, 1985; Polnikov, 1989, 1990, 1994, 2001; Resio and Perry, 1991)¹. Up to present, this part of the wind wave theory is practically solved (possibly excluding some details of long term nonlinear evolution (Lavrenov and Polnikov, 2001), and the main interest is addressed to implementation of the theory into practical numerical modeling.

Nowadays it is evident that the exact calculation of the kinetic integral can not be directly introduced into operational ocean wave models due to the large consuming time for this calculation. Therefore, one should use some kind of approximation to the exact integral. This is an important problem in the practice of wind wave numerical modeling, which, in turn, has theoretical and practical aspects. This paper is just devoted to consideration of the former. The latter will be considered in a separate paper.

The outline of the paper is the following. In Sect. 2 several present approximations of the two- dimensional NL-term are discussed, and main unsolved tasks are posed. Section 3 is devoted to a general consideration of the problem. From this consideration it follows that the most advanced approximations have the same mathematical root. A problem of approximation efficiency and optimization of approximation is posed and solved. In Sect. 4 a mathematical measure is introduced to estimate a relative error of any approximation with respect to exact 2-D nonlinear energy transfer among waves. The theory derived is used on the example of two alternative approximations in Sect. 5. In Sect. 6 two new versions of DIA are proposed, and their greater accuracy and efficiency are shown. Section 7 contains the final conclusions.

2 Statement of the problem

First of all, we should mention that there are several proposals dealing with numerical and analytical approximations for the kinetic integral. They have been presented both for one- dimensional (1-D) nonlinear transfer (Barnett, 1968; Resio, 1981; Zakharov and Smilga, 1981) and for the two-dimensional (2-D) one (Hasselmann et al, 1981; Hasselmann et al., 1985, Polnikov, 1991; Zakharov and Puchkarev, 1999, Lin and Perry, 1999; Hashimoto and Kawagushi, 2001; Van Vledder, 2001; and so on). For additional bibliography one may be referred to books: SWAM group (1985), Efimov and Polnikov (1991), Komen et al. (1994). In this paper we shall mainly focus on the approximations of the 2-D nonlinear transfer, which are more relevant for the modern wind wave modeling.

Among numerous 2-D approximations, one can select only a few that are theoretically well substantiated. The main feature of such an approximation should be its direct mathematical relation to the original kinetic integral. Actually, these approximations are as follow.

¹For additional bibliography, see Komen et al., 1994.

- 1. Diffusion approximation (DA) which was for the first time proposed in Hasselmann et al. (1981) and later elaborated in Zakharov and Pushkarev (1999) and Jenkins and Phillips (2001);
- 2. Discrete Interaction Approximation (DIA) proposed in Hasselmann et al. (1985) and elaborated in Hashimoto and Kawagushi (2001), Van Vledder (2001);
- 3. Reduced Integration Approximation (RIA) proposed in Lin and Perry (1999).

The dates of references mentioned permit us to hope that we consider the state-of-the art situation.

Here we shall try to analyze the main points of these approximations with the aim to find a way for an optimal solution of the problem. As far as a diffusion approximation proposed by Hasselmann (1981) is not practically used, we start our analysis from the more widely used Discrete Interaction Approximation.

2.1 Discrete interaction approximation

The main idea of DIA is to take into account only one certain configuration of the four interacting waves. To do this, Hasselmann et al. (1985) have proposed to use the configuration, which in the polar coordinates (σ , θ) has the form:

- (1) $k_1 = k_2 = k$, where the arbitrary wave vector k is represented by σ and θ ;
- (2) $\mathbf{k}_3 = \mathbf{k}_+$, where \mathbf{k}_+ is represented by $\sigma_+ = \sigma(1 + \lambda)$ and $\theta_+ = \theta + \Delta \theta_+$; (10a)
- (3) $\mathbf{k}_4 = \mathbf{k}_-$, where is represented by $\sigma_- = \sigma (1 \lambda)$ and $\theta_- = \theta - \Delta \theta_-$;

(4) In consistency with the conditions (5) and (6), parameters of the configuration are

$$\lambda = 0.25, \ \Delta \theta_+ = 11.5^\circ, \ \text{and} \ \Delta \theta_- = 33.6^\circ.$$
 (10b)

In such an approach, in accordance with Eq. (4), the nonlinear transfer at all mentioned k-points takes the form

$$\frac{\partial N(\boldsymbol{k}_{-})}{\partial_{t}} = I(\boldsymbol{k}, \, \boldsymbol{k}_{+}, \, \boldsymbol{k}_{-}), \, \frac{\partial N(\boldsymbol{k}_{+})}{\partial t} = I(\boldsymbol{k}, \, \boldsymbol{k}_{+}, \, \boldsymbol{k}_{-}),$$
$$\frac{\partial N(\boldsymbol{k})}{\partial t} = -2I(\boldsymbol{k}, \, \boldsymbol{k}_{+}, \, \boldsymbol{k}_{-}), \quad (11)$$

where

$$I(\mathbf{k}, \, \mathbf{k}_{+}, \, \mathbf{k}_{-}) = Cg^{-8}\sigma^{19} \left[N^{2}(\mathbf{k})(N(\mathbf{k}_{+}) + N(\mathbf{k}_{-})) - 2N(\mathbf{k})N(\mathbf{k}_{+})N(\mathbf{k}_{-}) \right]$$
(12)
$$D(\tilde{\Delta}\sigma, \, \tilde{\Delta}\theta).$$

In Eq. (12) the fitting constant is C = 3000, and $D(\tilde{\Delta}\sigma, \tilde{\Delta}\theta)$ is the differential expressed in non-dimensional increments of the integration grid, $\tilde{\Delta}\sigma$, $\tilde{\Delta}\theta$, (for the fixed grid, D is a constant). The net nonlinear transfer at any fixed (σ, θ) -point is found by the procedure of running of Eqs. (11) through all points of the frequency-angle integration grid $\{\sigma_i, \theta_i\}$.

The main advantage of this approximation is its evident simplicity and rather good efficiency for certain initial spectra (Hasselmann et al., 1985). For this reason, it is widely used in practical wave modeling. The so-called WAM model (Komen et al., 1994) is an example of a successful implementation of DIA. One of the technical shortages of DIA routine used in WAM is the presence of intermediate and cumbersome interpolation operations provided by the mismatch of spectral grid nodes and the vectors $\mathbf{k}_+, \mathbf{k}_-$. This leads to a large increase in the integration time.

In some papers (Polnikov, 1991; Zakharov and Pushkarev, 1999; Van Vledder, 2000; Lin and Perry, 1999; Jenkins and Phillips, 2001) it was mentioned that the accuracy of DIA is not reasonable for the JONSWAP spectrum. This stimulated some authors to a search for a more efficient approximation of the nonlinear term (see, for example, papers mentioned above and Hashimoto and Kawagushi, 2001; Van Vledder, 2001). Authors of the two latter papers tried to modify DIA by introduction of new and multiple configurations. However, it seems that these efforts did not significantly change the situation. Thus, the problem of DIA improvement is not achieved yet.

2.2 Diffusion approximation

For the first time the diffusion approximation (DA) was proposed in Hasselmann et al. (1981) by considering the exact integral (4) in the small scattering angle approximation. But the final expression for DA in this paper is rather complicated. A more elegant form of DA was proposed in Zakharov and Pushkarev (1999) and Jenkins and Phillips (2001) recently. Their results are not based on the explicit expression for the exact kinetic integral; they used the conservation laws (7)–(9) and inspection of some particular analytical solutions. Nevertheless, as it was shown in Polnikov (2002), that both of these proposals can be derived directly from the exact kinetic integral, if one estimates the final result for integral (4) as the contribution of the most contributive configuration

$$\boldsymbol{k}_1 \cong \boldsymbol{k}_2 \cong \boldsymbol{k}_3 \cong \boldsymbol{k}_4. \tag{13}$$

The use of this configuration is the main idea of the diffusion approximation. We note here that sometimes this idea is mentioned as the hypotheses of locality for nonlinear interactions.

Detailed analysis of the DA is given in Polnikov (2002). For this reason we will not elaborate here on this point but restrict ourselves to the following remarks.

First, the original expression by Hasselmann et al. (1981) for DA will not be considered here due to its relative complexity. Secondly, the most recent approximation by Jenkins So, the diffusion approximation due to Zakharov and Pushkarev (1999) (ZPA) has the following form

$$\frac{\partial N(\boldsymbol{k})}{\partial t} = \frac{c'}{\sigma^3} L\left(n^3(\boldsymbol{k})\sigma^{24}\right),\tag{14}$$

where L is the linear differential operator of the following form

$$L = \frac{1}{2} \frac{\partial^2}{\partial \sigma^2} + \frac{1}{\sigma^2} \frac{\partial^2}{\partial \theta^2}$$
(15)

and c' is the fitting constant of the order of 0.05 (Polnikov, 2002)².

It is important to note that according to Polnikov (2002), a preliminary (very qualitative) estimation of the relative accuracy of ZPA is of the order of 50%. Herewith, a rigorous mathematical definition of accuracy was not introduced in Polnikov (2002), and the real accuracy of ZPA (and DIA as well) is now known. Thus, we should state that the point of estimation of accuracy for the NL-term approximation is unsolved yet.

2.3 Reduced integration approximation

Another kind of approximation was proposed in Lin and Perry (1999), which was called as the Reduced Integration Approximation (RIA). To derive the latter, Lin and Perry (1999) placed their most attention on the fact that the integrand in Eq. (4) is growing infinitely for some configurations when $k_2 \rightarrow k_4$ and $k_1 \rightarrow k_3$ (the singularity of integrand mentioned earlier). It permitted them to reduce the threefold integral (4) to a quasi-linear one with a rather small interval of variations for k_1 and k_3 in the vicinity of k_4 . By this manner they reduce the time of integration radically with conserving a reasonable accuracy of the final result for the nonlinear transfer. In our notations, RIA takes the form

$$\frac{\partial n(\mathbf{k}_4)}{\partial t} = 4\pi \int_{k_4 - \Delta k}^{k_4 + \Delta k} dk_3 \int_{\theta_4 - \Delta \theta}^{\theta_4 + \Delta \theta} d\theta_3$$
$$M^3(\mathbf{k}_3, \, \mathbf{k}_4, \, \mathbf{k}_3, \, \mathbf{k}_4) J_\delta(\mathbf{k}_3, \, \mathbf{k}_4, \, \mathbf{k}_3, \, \mathbf{k}_4) N3(\mathbf{k}_3, \, \mathbf{k}_4), \tag{16}$$

where $J_{\delta}(\mathbf{k}_3, \mathbf{k}_4, \mathbf{k}_3, \mathbf{k}_4)$ is a rather complicated Jacobian due to integration of delta-functions in Eq. (4), $N3(\mathbf{k}_3, \mathbf{k}_4)$ is the proper cubic form of wave spectra, and Δk , $\Delta \theta$, are the special fitting parameters of the approximation.

As one can see, the net expression of the reduced integral in RIA is more complicated when compared to DIA and DA. In addition to this, RIA needs to use matrix elements M for calculations, which have to be calculated and stored previously. All these restrictions reduce the efficiency of RIA. Comparison of configurations used in Eq. (16) and Eq. (13) shows that RIA can be considered as a mixed version of DIA and DA. It looks like some a kind of multiple DIA but for DA-configuration (13). Thus, the properties of RIA resemble ones for DA (see Lin and Perry, 1999).

Finally, we should mention here again that neither Lin and Perry (1999) nor others authors proposed a rigorous mathematical procedure for inter-comparison of different approximations for the exact kinetic integral. For this reason we have not ways to get a rigorous estimating the accuracy of any certain approximation.

All these uncertainties gave rise to the appearance of new approximations among which different modifications of DIA (Hashimoto and Kawagushi, 2001; Van Vledder, 2001) and DA (Zakharov and Pushkarev, 1999; Jenkins and Phillips, 2001; Polnikov, 2002) are looked like as the most promising.

2.4 Problems posed

Now we can formulate a set of tasks of the present paper, concerning unsolved theoretical points of the problem for optimization of the *NL*-term approximation. They are as follows. 1) To consider the problem in general form and to find the common and differing features of the approximations mentioned above. 2) To establish a rigorous measure of efficiency of approximations for the exact kinetic integral. 3) To define a mathematical procedure for comparing accuracy and efficiency of different approximations. 4) To test DIA and ZPA with respect to their accuracy and efficiency. 5) To propose an approximation, which is more efficient than any other considered above. All these problems will be treated and solved in the present paper.

3 Phylosophy of the approach

3.1 General point of view

We start from simplifying a representation of the kinetic integral. According to Polnikov (1989), after an exact analytical integration of the wave vector delta-function with respect to k_2 and the frequency delta-function with respect to θ_1 , in polar coordinates the integral (4) takes the form

$$\frac{\partial N_4}{\partial t} = c \sum_{\pm} \int \int \int M_{1\pm,\,2\pm,\,3,\,4}^2 \frac{J}{k_{2\pm}} \\ \left(N_{1\pm} N_{2\pm} (N_4 + N_3) - N_4 N_3 (N_{1\pm} + N_{2\pm}) \right)$$
(17)
$$d\sigma_1 d\sigma_3 d\theta_3,$$

where *c* is the dimensional coefficient depending on *g*, $N_i = N(\sigma_i, \theta_i)$, *J* is the Jacobian of the frequency delta-function integration, $k_{2\pm}$ is the modified wave number corresponding to the point $(\sigma_2, \theta_2^{\pm})$, and the index (\pm) represents a mirror symmetry configuration with respect to referring vector $k_a = k_3 + k_4$ (for details, see Polnikov, 1989). An explicit

 $^{^{2}}$ The optimal fitting constants both in DIA and DA are varying in dependence of the spectral shape.



Fig. 1. Visual representation of the configurations permitted by Eqs. (5) and (6). Contour lines correspond to the possible end points of interacting vectors. Example is shown for the original DIA configuration given by Eq. (10) (Hasselmann et al., 1985).

form of J is not needed for understanding the after going text. We only note that the Jacobian detects a 3D singular surface, which defines the most contributive configurations of interacting waves.

From Eq. (17) one can see that in a discrete representation of the integration grid $\{\sigma_{ig}, \theta_{jg}\}$, the integral (17) can be written as a simple sum of the form

$$\frac{\partial N_4}{\partial t} = c \sum_i R_i (N3)_i D_i = \sum_i B_i \left(N_{1\pm} N_{2\pm} (N_4 + N_3) - N_4 N_3 (N_{1\pm} + N_{2\pm}) \right)_i, \quad (18)$$

where *i* is the number of allowed interacting configuration. In Eq. (18) R_i is the spectrum independent part of the integrand in Eq. (17), $(N3)_i$ is the cubic spectral form, and D_i is the integration differential in the frequency-angle space (the sub-index *i* means that the values are taken for a proper interacting configuration). After grouping of all spectrum independent parts of the summands, $cR_iD_i \equiv B_i$, the final sum takes a simple form. Note that B_i is the fixed constant for a certain configuration of four vectors k_1 , k_2 , k_3 , k_4 . In the (σ, θ) -space B_i is proportional to the scaling factor σ^{19} (see Eq. 12) and is depending on the angles between vectors of configuration only. ³

Analysis of Eq. (18) allows us to find common and differing features of various approximations mentioned in Sect. 2 and to generate ideas about optimal approximation for the integral Eq. (17).

As we mentioned above, configurations of four wave vectors k_1 , k_2 , k_3 , k_4 contributing to the integral should meet the resonant conditions Eq. (5) and Eq. (6). Herewith, the

most contributive configurations include the points at the singular surface. As was shown in Masuda (1980), Polnikov(1989), some of such configurations are corresponding to the condition

$$k_1 = k_2 = (k_3 + k_4)/2 = k_a/2 \tag{19}$$

where k_a is the reference vector of the configuration. Comparing Eq. (19) with Eq. (10) one can see that this condition is the basis of DIA.

It is well known that a lot of other configurations are located at the singular surface. They can be represented by the "figure-of-eight" in the Longett-Higgins diagram (Fig. 1). The most contributive of them (for which the value of R_i is going to infinity with the greatest rate) corresponds to the configuration

$$k_1 = k_2 = k_3 = k_4. \tag{20}$$

According to Eq. (13), this is just the configuration used in DA (and RIA).

Thus, from this consideration one can see that DIA, DA and RIA has the same root. All of them are based on the account for contribution of certain configurations at the singular surface. In fact, original DIA and different variants of DA are based on a one configuration only. The multiple DIA (with its variations) or RIA has a number of configurations and does not essentially differ from DIA. Thus, the approximations under consideration have a common origin.

The only technical difference between DA and DIA (RIA) is provided by the fact that the cubic spectral form N3 becomes equal to zero for configuration (20). For this reason, one should take a Tailor's expansion of the cubic form in the vicinity of the point (20), which leads to the diffusion operator representation of the DA, as was shown in Polnikov (2002).

From the analysis above, we can conclude:

- 1. The most adequate (theoretically grounded) approximations employ the wave-number configurations allocated at the singular surface of the integrand in Eq. (17).
- 2. The only difference between the approximations DIA, DA, and RIA is the choice of set of the contributing configurations.
- 3. The optimal approximation for integral (17) can be defined as one leading to the highest efficiency of approximation. Such a kind approximation can be constructed by means of choosing a special set of the contributing configurations, if the definition of efficiency is given.

3.2 Efficiency of approximation

The term "efficiency of approximation" is often used in the literature mentioned above, but it was never formulated in exact mathematical form. Here we address this point.

Each approximation has its own accuracy and speed of computation. Therefore, it is useful to introduce the following two quantitative parameters: the relative error, ε_{rel} , and

³There is an additional (rather strong) dependence of B_i on its location at the singular surface (see Polnikov, 1989), but it is less important from the configuration point of view.

the time of one-step calculation, τ .⁴ These two parameters should be used in the definition of efficiency of approximation. Usually these two parameters are in a competitive context: one is working against the other. In general, it is not so simple to say which of them is more preferable. To do this, one needs some postulates depending on the specific goal. For this reason, it seems that the term "efficiency" is the more intuitive (or qualitative) parameter of the approximation. Nevertheless, in this sub-section we introduce some postulates and formulate the meaning of the term "efficiency" in a mathematical form.⁵

Firstly, we state here that the accuracy (or relative error) of the numerical approximation of the integral (17) is more important than other aspects. Let us consider this point in more detail.

To calculate the relative error, ε_{rel} , of the approximation, one should have a reference value of the exact kinetic integral. As far as the integral (17) in a closed form is not known, the so-called "exact calculation" is to be done numerically as well. This procedure has its own relative error with respect to the "exact theoretical value" of the kinetic integral (which corresponds to the theoretical definition of integration)⁶. The relative error of the "exact calculation" can be estimated by traditional means, i.e. by making the integration grid resolution to be finer and finer. Eventually, we can state that the "exact calculation" is executed by using a proper mathematical algorithm and by using a certain frequency-angle grid for integration in Eq. (17).

Appropriate algorithms for this task are well known (se references above). How to choose the standard frequencyangle grid, and how should be the relative error of "the exact calculation"? To answer these questions, one should "a priori" introduce the lower limit value of the relative error for the approximated calculation of NL-term, ε_{lim}^{ap} . Taking in mind numerous ambiguities of the source term in Eq. (1), one may postulate that the lower limit of relative error of the approximation should be not smaller than 10–15%, i.e. $\varepsilon_{lim}^{ap} \approx 0.10 - 0.15$. This means that for practical aims there is sufficient to have an approximation for NL-term the relative error of which meets the following condition

$$\varepsilon_{rel} \ge \varepsilon_{lim}^{ap} \cong 0.10 - 0.15. \tag{21}$$

On the other hand, it is reasonable to assume that the requested relative error of the "exact calculation", ε_{lim}^{ex} , must be at least one order smaller:

$$\varepsilon_{lim}^{ex} \cong 0.1 \ \varepsilon_{lim}^{ap} \cong 0.01 - 0.015.$$

⁴The mathematical definition of ε_{rel} will be given later in Sect. 4. Relative accuracy, α_{rel} , is related to the relative error, ε_{rel} , by the evident simple ratio: $\alpha_{rel} = 1 - \varepsilon_{rel}$. In the after going text both terms are used.

⁵Here we consider the efficiency for the one-step calculation of the kinetic integral. The point of the long-term efficiency is rather similar but more complicated due to the absence of the commonly recognized opinion about features of the long-term solution of the kinetic Eq. (4) (see discussion in Lavrenov and Polnikov, 2001).

⁶Note that in the aspect of accuracy, the time of "exact calculation" does not play any role. From this, one may conclude that for estimation of the relative error of the *NL*-term approximation, the value of ε_{lim}^{ex} should be of the order of 1–2%. Thus, the standard frequency-angle grid may be of any kind which provides the necessary ε_{lim}^{ex} . This is the first theoretical conclusion dealing with the accuracy consideration. As seen, it leads to the constraint (22) on the features of the standard integration grid needed for the estimation of ε_{rel} .

Existence of different standard frequency-angle grids is known from numerous calculations (Hasselmann et al., 1981; Masuda, 1980; Polnikov, 1989; and so on). A certain example for one of them will be given later in the following section.

Consider now the issue of computation time. For practical purposes, one may separate the following two types of computation time for the *NL*-term calculation: (1) a "relative integrating time", and (2) a "one-step time". The fist type of time is defined by the relative part of the CPU- time, *RP*, taken by the *NL*-term sub-routine in numerical model calculations as a whole. The second type of time is defined by the real CPU-time, τ , needed for the calculation of *NL*-term at one step in time.

A definition of efficiency is different for different values of RP. This fact results in to possibility of introduction the following two meanings: the first type efficiency, Eff_1 , and the second type efficiency, Eff_2 .

In the first case, when $RP \ll 1$, there is a natural upper limit for "the relative integrating time" of the *NL*-term calculation, based on the following practical consideration. If the time of *NL*-term calculations takes less than 15–20% of the total time of prognostic calculations in a numerical model $(RP \le 0.15 - 0.2)$, there is no need to reduce *RP* more, as far as the other part of the model takes the major time of CPU. Therefore, in such a case, the first type efficiency may be defined by the relative error of *NL*-term approximation only. A heuristic formula for efficiency could be proposed in the following form

$$Eff_1 \propto (\varepsilon_{rel})^{-p}.$$
(23)

Here the power p is introduced to emphasize the role of accuracy (or relative error) in the definition of approximation efficiency. As far as in the "error theory" just the second power of error plays a main role (for example, for summation of errors), we may specify the power value p to be equal to 2. Then, if we accept the highest limit of efficiency to be about 100 units, the best formula for Eff_1 corresponding to the lower limit of ε_{rel} (21) takes the form

$$Eff_1 = (\varepsilon_{rel})^{-2} \tag{24}$$

In the second case, when the relative part of *NL*-term calculations is sufficiently large (for example, RP > 0.2), it is reasonable to suppose that the efficiency of approximation depends inversely on the one-step time of *NL*-term calculation, τ , to the power of *RP*. In accordance with Eq. (24) Eff_2 can be given as

$$Eff_2 = (\varepsilon_{rel})^{-2} (T_{ref}/\tau)^{RP}$$
(25)

where T_{ref} is the reference time introduced for normalization of the value of the time τ . Here, T_{ref} should be always smaller than τ (see the definition of T_{ref} below). The power RP is introduced to emphasize the role of calculation time as an auxiliary parameter. It means that for rather small values of RP (when RP becomes much less than 1 and τ is of the order of T_{ref}), the second type efficiency, Eff_2 , should degenerate into the efficiency of the first type, Eff_1 .

To close the point, we need to specify the reference time T_{ref} . To do this, we accept the definition that T_{ref} is the time taken by the calculation of *NL*-term in a "fast one-configuration approximation" given by Eqs. (11) and (12) without any intermediate interpolation procedures ⁷. In such a case, the more configurations and more intermediate (interpolation) procedures are involved into a certain approximation, the greater a real one-step time τ is for this approximation. Consequently, the ratio T_{ref}/τ becomes smaller. As a result, the total efficiency is balanced by the relative error of the approximation according to Eq. (25).

Final conclusions of this sub-section are the following.

- 1. The term "efficiency" is rather an intuitive and qualitative parameter of the approximation.
- 2. The efficiency of a numerical approximation of the *NL*-term depends on several parameters in the calculation process. They are: the relative error of approximation, ε_{rel} ; time of one-step integration, τ ; relative part of time taken by *NL*term subroutine from the whole time of prognostic calculations, *RP*; and some heuristic magnitudes, such as the power *p* in Eq. (23) and reference time of "the simple one-configuration approximation", T_{ref} .
- 3. The role of accuracy is more important than the role of the computation time in the approximated calculation of the *NL*-term. It means that for approximations with equivalent efficiency, the one with more accuracy (or less error ε_{rel}) is preferable.
- 4. The efficiency of approximation has its natural upper value (of the order of 100 conventional units) which is absolutely sufficient for practical goals.
- 3.3 Optimization of approximation

In accordance with the said above, the efficiency of any approximation is balanced by two competitive characteristics: the accuracy and the computation time. By definition, the optimal approximation corresponds to the best balance between relative error, ε_{rel} , and the one-step time of calculation, τ , which provides the highest efficiency. Thus, an optimal approximation for integral (17) is provided by the choice of optimal set of the contributing configurations.

Here we should note that for any fixed set of configurations, the value of the relative error of approximation, ε_{rel} , depends on the shape of spectrum under the integral. For this reason an estimation of the efficiency is not so simple as it seems from a first sight. We believe that the finding of the efficiency of approximation is not a purely technical task but, rather, is some kind of an expertise process. To carry out this process, one needs to select an appropriate set of reference spectra. Such a kind of set will be proposed in Sect. 5 after getting a preliminary experience of efficiency estimating. In the course of these preliminary calculations a certain specification will be done for the choice of the reference time of calculation, T_{ref} .

4 Measure of relative error

We consider the following two-dimensional functions: (a) The exact nonlinear energy transfer in the frequency-angle space, $T_{ex}(\sigma, \theta)$, obtained by a certain numerical algorithm at the standard integrating grid for a certain spectral shape of the energy spectrum $S(\sigma, \theta)^{8}$; (b) The approximated nonlinear energy transfer $T_{ap}(\sigma, \theta)$, at the same space, obtained for the same energy spectrum and at the same frequency-angle grid by any numerical algorithm.

The problem to be addressed is how to estimate the nondimensional relative error of the approximation, ε_{rel} rigorously and systematically. Regarding the 2-D nonlinear transfer, this problem was never discussed in the literature. Therefore, some of the following statements will be rather heuristic.

Some initial specifications are now opportune⁹.

- 1. The 2-D functions $T_{ex}(f, \theta)$ and $T_{ap}(f, \theta)$ are to be given at a properly chosen frequency-angle grid (f_{ig}, θ_{jg}) (at the so-called "standard integrating grid" providing a necessary accuracy of $T_{ex}(f, \theta)$, see Sect. 3).
- 2. For simplifying the comparison of the results for different spectral shapes, all values of $T_{ex}(f, \theta)$ and $T_{ap}(f, \theta)$ are to be normalized by a certain dimensional coefficient depending on the peak frequency, f_p , and the peak value of the 2-D spectrum, S_p . In our calculations the nonlinear transfer functions are calculated in conventional units of the normalizing coefficient, C_n , introduced in Polnikov (1989)

$$C_n = (\pi/16)g^{-4}f_p^{11}S_p^3.$$
(26)

3. Due to an ambiguity in the fitting coefficient for the approximated transfer, $T_{ap}(f, \theta)$ (see Sect. 2), one should adjust the latter to $T_{ex}(f, \theta)$ by an additional adjusting coefficient, C_{ad} . In our work this coefficient is estimated by means of the least-square-root method in ac-

⁷A point of the spectrum interpolation is discussed in Sect. 6.

⁸The meaning of the terms "exact" and "the standard integrating grid" were clarified earlier in Sect. 3.

⁹In the after going text we prefer to use the cyclic frequency $f = \sigma/2\pi$ instead of angular one, σ .

cordance with the condition

$$\int_{\Omega} \left(T_{ex}(f,\,\theta) - C_{ad}^{(ia)} T_{ap}^{(ia)}(f,\,\theta) \right)^2 df d\theta = \min.,\,(27)$$

where (ia) means the individual index of approximation, and Ω is the fixed part of the frequency-angle space used for estimation of C_{ad} . In our calculations the domain Ω covers the whole frequency-angle grid under consideration. With some algebra, one can find the following expression for C_{ad} :

$$C_{ad}^{(ia)}(\Omega) = \frac{\int \Omega}{\Omega} \frac{T_{ex}(f,\,\theta) T_{ap}^{(ia)}(f,\,\theta) df d\theta}{\int \Omega \left(T_{ap}^{(ia)}(f,\,\theta)\right)^2 df d\theta}.$$
 (28)

4. Finally, the properly adjusted nonlinear transfer for the *(ia)*-th approximation, used for the error estimation, has the following value

$$T_{ap}^{(ia)}(f,\,\theta) = T_{ap,\,init}^{(ia)}(f,\,\theta)^* C_{ad}^{(ia)},$$
(29)

where the domain factor Ω is omitted for simplicity.

Now one can introduce a formula for the relative error, ε_{rel} . For this aim, a traditional measure can be used. For example,

$$\varepsilon_{rel}^{(ia)}(\Omega_{\varepsilon}) = \left(\frac{\int\limits_{\Omega_{\varepsilon}} \left(T_{ex}(f,\,\theta) - T_{ap}^{(ia)}(f,\,\theta)\right)^2 df d\theta}{\int\limits_{\Omega} \left(T_{ap}^{(ia)}(f,\,\theta)\right)^2 df d\theta}\right)^{1/2} (30)$$

where Ω_{ε} is the fixed part of the frequency-angle space used for estimation of ε_{rel} . However, the measure (30) is too smooth, as far as it is not sensitive to the location of the points where the nonlinear transfer changes the sign. Nevertheless, it seems reasonable that the proper measure must include in its definition this very important feature of the *NL*-transfer responsible for the spectrum shape evolution. For this reason we prefer the formula which is more sensitive to this feature of the nonlinear transfer function. After some analysis we made the option for the following measure of ε_{rel} :

$$\varepsilon_{rel}^{(ia)}(\Omega_{\varepsilon}) = \left(\frac{\int\limits_{\Omega_{\varepsilon}} \left|\frac{T_{ex}(f,\theta) - T_{ap}^{(ia)}(f,\theta)}{T_{ex}(f,\theta)}\right|^{m} df d\theta}{\int\limits_{\Omega_{\varepsilon}} df d\theta}\right)^{1/m}$$
(31)

with the choice of m = 1. The more typical choice, m = 2, was not acceptable because it gives the relative error measure too sensitive to the zero-crossing feature of the nonlinear transfer.

We should note that the definition (31) could be easily transformed to the case of the 1-D NL-transfer, T(f) =

 $\oint T(f, \theta)d\theta$. In fact, for the 1-D relative error, $\varepsilon_{rel, 1-D}$, the transformation of Eq. (31) is

$$\varepsilon_{rel, 1-D}^{(ia)}(\Omega_{\varepsilon}) = \frac{\int \limits_{\Omega_{\varepsilon}} \left| \frac{T_{ex}(f) - T_{ap}^{(ia)}(f)}{T_{ex}(f)} \right| df}{\int \limits_{\Omega_{\varepsilon}} df}$$
(32)

In principle, the 1-D relative error is less interesting in our analysis. Nevertheless, in the present paper, it will be considered in some cases, for generality.

Finally, we need to say some words about a choice of the domain Ω_{ε} . In the general case, it is possible to analyze relative errors for several types of domains. But later we shall mainly deal with results only for the 10%-threshold domain defined by the following ratio

$$\Omega_{\varepsilon} = \Omega_{10\%} \in |T_{ex}(f, \theta)| \ge 0.1R, \tag{33}$$

where

$$R = T^+ - T^-. (34)$$

Here T^+ is the positive extremum of the exact 2-D nonlinear transfer and T^- is the negative one. Owing to the integral feature of the relative error (31), hereafter it is called the mean relative error (MRE). For generality, the first estimations of errors will be presented for 20%-threshold domain as well. In the following section the proposed approach will be applied to estimation of the MRE and efficiency for two alternative approximations: DIA and ZPA, the most prospective ones at present time.

5 Estimations of errors and efficiency for DIA and ZPA

5.1 Numerical specifications

For the calculation of the exact nonlinear energy transfer, $T_{ex}(\sigma, \theta)$, we have used the algorithm proposed in Polnikov (1989) and the code by the same co-author. The code for DIA was extracted from the widely used version of WAM (Cycle 4) with minor programming adjustments to our task. The code for ZPA was written in accordance with Eqs. (14) and (15) (for details, see Polnikov, 2002).

The standard frequency-angle grid, (f_i, θ_j) , used for our calculations, is defined by the following ratios

$$f(i) = f_0 \cdot e^{i-1},$$
(35)

$$\theta(j) = -\pi + j \cdot (\pi/18), \tag{36}$$

with the following values of the grid parameters: $f_0 = 0.7462$, e = 1.05, and $1 \le i \le 41$, $0 \le j \le 35$. The relative error of "the exact estimations" for $T_{ex}(\sigma, \theta)$ is about 1–3%.

For more generality we have used the following two-mode spectrum representation

$$S(f, \theta) = S_1(f, \theta, f_{p1}, \theta_{p1}, \gamma_1, s_1) + R2 \cdot S_2(f, \theta, f_{p2}, \theta_{p2}, \gamma_{21}, s_2),$$
(37)



Fig. 2. Visual representation of 1-D-nonlinear transfer $\partial S(f)/\partial t = T_s(f)$ for the exact calculation and adjusted DIA and ZPA estimations for the run 1.

where each of modes has a typical *JONSWAP* spectrum of the form

$$S(f, \theta, f_p, \theta_p, \gamma, s) = \alpha f^{-5} \exp(-1.25(f_p/f)^4)$$
$$\frac{\exp[-(f - f_p)^2 / 0.01 f_p^2]}{\gamma_J} \Psi(s, \theta, \theta_p).$$
(38)

In Eq. (38) the coefficient α is taken equal to 1, and the angular spreading function is of the form

$$\Psi(s,\,\theta,\,\theta_p) = I_s \cos^s(\theta - \theta_p) \tag{39}$$

with normalization coefficient I_s taken equal to 1, for simplicity (as far as we use for comparison the normalized values of the nonlinear transfer, see Sect. 4). Coefficient *R*2 is responsible for the relative intensities of the modes. The extended set of parameters defining the set of the spectrum used in our investigations is presented in Table 1.¹⁰ On the basis of this set, after getting some experience, the standard set of spectrum is postulated for the further studies.

5.2 MRE estimations and their analysis

Results of the mean relative error, ε_{rel} , for DIA and ZPA are presented in Tables 2 and 3. For generality of consideration, at the this stage of MRE estimations we represent both the 10%-threshold error, MRE (10%), and 20%-threshold error, MRE (20%) (see definition in Sect. 4). Herewith, the most attention is paid to the MRE (10%), as this is the more appropriate error parameter from the practical point of view.

To represent the kind of errors we did visualize the 1-D-Nl transfer for both approximations and for two runs of spectral

shapes (Figs. 2 and 3). But we should mention that this is rather a qualitative representation of MRE, whilst the tables give the quantitative one. For this reason we will not dwell on these pictures below.

Before considering efficiency, let us analyze the results for the MRE. Such kind of estimations is presented in literature for the first time, and, for this reason, they have their own interest.

5.2.1 Discrete interaction approximation

From Table 2 one can see the following:

- The 10%-threshold mean relative error, MRE (10%), is, as a rule, greater than 20%-threshold mean relative error, MRE (20%) (as it may be expected from theoretical consideration). But there are five cases when the opposite situations take place: runs 3, 10, 13–15.
- Variation of values for MRE (10%) due to dependence on the spectrum shape is important:

$$22\% \le MRE (10\%) \le 69\%$$
 for 2-D NL-transfer,
 $22\% \le MRE (10\%) \le 108\%$ for 1-D NL-transfer

- Variation of values for the adjusting coefficient C_{ad} due to dependence on the spectrum shape is remarkable: $1.13 \le C_{ad} \le 2.81$.

The detailed analysis of MRE (10%) for all the runs allows us to draw the following conclusions.

1. The error values depend essentially on the spectral shape parameters.

¹⁰A justification of the choice for the spectral shapes under consideration one can find in Polnikov (1989).

No.	f_{p1} ,	$oldsymbol{ heta}_{p1}$,	γ_1	S_1	R2	f_{p2} ,	$\theta_{_{p2}},$	γ_2	s_2
of run	conv.un	degrees				conv.un.	degrees		
1	1	0	1	2	0				
2	1	0	1	8	0				
3	1	0	3.3	2	0				
4	1	0	3.3	12	0				
5	1	0	1	8	0.4	2	0	3.3	4
6	1	0	1	8	1.2	2	0	3.3	4
7	1	0	1	8	1.2	2	-60	3.3	4
8	1	0	1	8	0.4	2	-60	3.3	4
9	1	0	1	8	1.2	2	-60	3.3	8
10	1	0	1	8	1.2	2	-180	3.3	4
11	1	0	1	8	1	1	-80	1	8
12	1	0	1	8	1	1	-180	1	8
13	1(swell)	0	1	8	1.2	2	0	3.3	4
14	1(swell)	0	1	8	0.4	2	0	3.3	4
15	1(swell)	0	3	8	1.2	2	0	3.3	4
16	1(swell)	0	3	8	0.4	2	0	3.3	4

Table 1. A set of parameters for spectra used in calculations

Note: "(swell)" in the first column means that the power of spectrum tail fall is $S_1(\sigma) \propto \sigma^{-10}$.



Fig. 3. The same as in Fig. 2 but for the run 3.

2. The greatest errors take place for very narrow (in frequency and in angular spreading) spectra (runs 3, 4) which are typical in developing waves. This effect is proved by a special consideration of 4 cases with swell second mode (runs 13–16) (mixed sea cases). In the latter cases, MRE (10%) for DIA can reach 100% (for 1-D *NL*-transfer).

3. In principle, the mean ("ensemble") error is an artificial parameter. Nevertheless, this "expert-type" parameter

Table 2. MRE for 2-D-Nl and 1-D-Nl transfers in DIA

No. of	Type of	Adjusting	MRE(10%),	MRE(20%),
run	NL-transfer	coefficient C_{ad}	percents	percents
1	2D-Nl	1.37	22.0	0.166
	1D- <i>Nl</i>		19.7	
2	2D- <i>Nl</i>	1.25	35.9	20.9
	1D- <i>Nl</i>		21.7	
3	2D- <i>Nl</i>	1.33	68.9	78.6
	1D- <i>Nl</i>		70.2	
4	2D- <i>Nl</i>	2.44	73.3	56.1
	1D- <i>Nl</i>		69.8	
5	2D-Nl	1.21	36.2	24.3
	1D- <i>Nl</i>		82.7	
6	2D-Nl	1.13	41.6	36.1
	1D- <i>Nl</i>		71.8	
7	2D-Nl	1.31	46.7	30.2
	1D- <i>Nl</i>		57.4	
8	2D- <i>Nl</i>	1.52	35.2	24.2
	1D- <i>Nl</i>		25.3	
9	2D-Nl	1.17	49.0	30.0
	1D- <i>Nl</i>		45.5	
10	2D-Nl	1.49	43.4	44.6
	1D- <i>Nl</i>		76.2	
11	2D- <i>Nl</i>	1.44	40.3	31.6
	1D- <i>Nl</i>		29.3	
12	2D- <i>Nl</i>	1.69	24.4	15.9
	1D- <i>Nl</i>		51.1	
13	2D- <i>Nl</i>	1.32	64.5	69.9
	1D- <i>Nl</i>		92.7	
14	2D- <i>Nl</i>	1.35	51.5	60.2
	1D- <i>Nl</i>		90.7	
15	2D- <i>Nl</i>	1.33	64.7	69.9
	1D- <i>Nl</i>		93.7	
16	2D- <i>Nl</i>	2.81	68.1	46.4
	1D- <i>Nl</i>		108	

can be introduced for the practical aims as a mean value of MRE for the representative set of spectrum shapes.

On the basis of these conclusions we can propose the following recommendations.

- 1. The most representative set of spectrum shapes, which can be proposed for the next elaboration of DIA, should include runs 1, 2 and runs 3–7, 15, 16. The former two are considered as cases typical for the spectrum of developed wind waves, while the latter are cases where the spectrum shape is of developing and mixed sea type.
- 2. With respect to DIA efficiency, the most appropriate value of the relative error provided by DIA for 2-D-*Nl* can be accepted as a simple averaging of MRE (10%) through the 9 representative cases mentioned above. The "expert" estimation is

$$\varepsilon_{rel}^{(DIA)} \cong 0.5. \tag{40}$$

- 3. Regarding DIA implementation, the "expert" value of C_{ad} is of the order of 1.6, which is in a good agreement with recommendations of the WAM group (Komen et al., 1994).
- 5.2.2 Zakharov-Puchkarev's diffusion approximation

From Table 3 one can see the following:



Table 3. MRE for 2-D-Nl and 1-D-Nl transfers in ZPA

No. of	Type of	Adjusting	MRE(10%),	MRE(20%),
run	NL-transfer	coefficient C_{ad}	percents	percents
1	2D-Nl	0.262	75.9	61.6
	1D- <i>Nl</i>		80.9	
2	2D- <i>Nl</i>	0.115	66.2	58.7
	1D- <i>Nl</i>		70.3	
3	2D- <i>Nl</i>	0.0083	103	88.1
	1D- <i>Nl</i>		106	
4	2D-Nl	0.0085	91.8	84.7
	1D- <i>Nl</i>		90.3	
5	2D-Nl	0.0580	74.8	77.3
	1D- <i>Nl</i>		78.1	
6	2D- <i>Nl</i>	0.022	86.5	86.1
	1D- <i>Nl</i>		75.3	
7	2D-Nl	0.0344	83.6	83.2
	1D- <i>Nl</i>		90.1	
8	2D-Nl	0.108	66.9	62.2
	1D- <i>Nl</i>		75.7	
9	2D-Nl	0.031	84.8	84.9
	1D- <i>Nl</i>		92.0	
10	2D-Nl	0.023	94.0	85.7
	1D- <i>Nl</i>		104	
11	2D- <i>Nl</i>	0.131	75.9	71.5
	1D- <i>Nl</i>		88.4	
12	2D-Nl	0.153	64.8	60.4
	1D- <i>Nl</i>		76.8	
13	2D- <i>Nl</i>	0.0089	99.5	80.2
	1D- <i>Nl</i>		103	
14	2D-Nl	0.0164	89.7	71.2
	1D- <i>Nl</i>		99.9	
15	2D-Nl	0.0092	100	79.8
	1D- <i>Nl</i>		105	
16	2D- <i>Nl</i>	0.0111	77.1	77.6
	1D- <i>NI</i>		73.1	

Fig. 4. Comparative diagram of relative errors for the original DIA and ZPA. In horizontal axes the number of run from Table 1 is presented.

- The relative errors for ZPA are greater than errors for DIA (in 1.5–2 times approximately), except of runs 5 and 10 for 1-D *NL*-transfer. A comparative diagram of errors for these two cases is presented in Fig. 4.
- The 10%-threshold mean relative error, MRE (10%), is, mainly, grater than 20%-threshold mean relative error, MRE (20%). There are only 2 cases when the opposite

situations take place: cases 5 and 9.

- Variation of values for MRE (10%) due to dependence on the spectrum is rather remarkable:
- $-64\% \le MRE (10\%) \le 103\%$ for 2-D *NL*-transfer,
- $-70\% \le MRE (10\%) \le 106\%$ for 1-D *NL*-transfer.
- Variation of values for the adjusting coefficient C_{ad} due to dependence on the spectrum is very essential: 0.26 $\leq C_{ad} \leq 0.0083$.

On the basis of the presented results for ZPA, the following conclusions may be drawn.

- 1. The relative errors for ZPA are greater than errors for DIA in 1.5–2 times approximately, and their values depend strongly on the spectral shape (Fig. 4).
- 2. The adjusting coefficient depends very strongly on the shape of the spectrum under consideration. It seems that there is not any reasonable coefficient C_{ad} , which can be used as the appropriate one for all the spectral shapes in this approximation.
- 3. From the point of view of accuracy and stability of values for C_{ad} , ZPA is less preferable than DIA.

Despite of the last conclusion, some practical recommendations can be given.

1. With respect to the ZPA efficiency, the "expert" value of the relative error for 2-D *NL*-transfer (averaged through the 9 representative runs) can be accepted of the order of 85%, i.e.

$$\varepsilon_{rel}^{(ZPA)} \cong 0.85 \tag{41}$$

2. In the case where ZPA is needed for practical use in numerical modeling, the most acceptable value for C_{ad} could be of the order of 0.05, according to recommendation given in Polnikov (2002).

5.3 Estimations of efficiency for DIA and ZPA

Due to the relatively long time of the *NL*-term calculations in DIA, the second type of efficiency given by Eq. (25) is more relevant to our analysis. To estimate efficiency, in addition to the estimation of relative error, we need proper estimations of the non-dimensional parameters $t_{rel} \equiv (T_{ref}/\tau)$ and *RP*. We have used a COMPAQ workstation for *NL*-term calculations and supercomputer NEC SX-4 for WAM (Cycle 4) application. The former was used for estimation of t_{rel} by the procedure SECNDS, while the latter was used for estimation of *RP* by the software PROFILE.

The procedure of efficiency estimation and results are as follows.

Table 4. Time parameters for DIA and ZPA

Approximation	τ_{10000} , s	10000 <i>T_{ref}</i> , s	t_{rel}	
DIA	2.26	0.31	0.14	
ZPA	0.36	0.31	0.86	

In fact, the following three codes are used in the efficiency estimation procedure. (1) A specially prepared oneconfiguration DIA adjusted to the integrating grid ("fast DIA") is used for estimation of T_{ref} ; (2) An original DIA (or ZPA) approximation is used for estimation of the proper value of τ ; (3) The WAM (Cycle 4) is used for estimation of *RP*.

To estimate T_{ref} , we have excluded the interpolation procedures and the functional derivative part from the original DIA (fast DIA)¹¹. After this we have estimated the CPUtime for 10000 cycles of the fast DIA. This time is just an estimation of the value 10000 T_{ref} .

Real one-step time, τ , for DIA and ZPA is estimated by the same "10000-step procedure" but for the original DIA-code (without the functional derivative part) and for our own ZPA-code. By this way we estimate an auxiliary value, $\tau_{10000} = 10000\tau$. Finally we have found the estimations presented in Table 4.

The relative part *RP* is estimated by the use of the PROFILE-procedure applied to the WAM (Cycle 4). For the total original DIA, the real value of $RP^{(DIA)}$ is of the order of 0.45. For the present calculations of the *NL*-term only (without functional derivative), one may estimate the value of $RP^{(DIA)}$ to be equal to 1/3 which is properly less that one known for the total DIA. Thus, the rest part of WAM takes about 65% of the whole time of wave forecasting. Note that this part of calculation time is not changed while changing *NL*-term approximation. Taking in mind the ratio $\tau_{10000}^{(DIA)}/\tau_{10000}^{(ZPA)} \cong 6$ following from Table 4, it is easy to calculate that the proper value of $RP^{(ZPA)}$ is of the order of 1/13. So, we have all needed parameters for estimations of efficiency for the approximations under consideration.

According to Eqs. (25), (40), and (41), the efficiency for the DIA and ZPA approximations is

$$Eff_2^{(DIA)} = (1/0.5)^2 (0.14)^{1/3} \cong 2;$$
 (42)

$$Eff_2^{(ZPA)} = (1/0.85)^2 (0.86)^{1/13} \cong 1.4.$$
 (43)

Thus, despite the relatively high speed of ZPA calculation of the NL-term, the real efficiency of DIA is greater due to its better accuracy.

Taking in mind that the limiting efficiency of approximation is of the order of 100 units, it seems that one has large "space" for construction of better approximations for NLterm. According to the conclusions of optimal approximations made in Sect. 3, the way of developing a new approxi-

¹¹About fast one-configuration DIA adjusted to the integration grid see below Sect. 6.

mation of the NL-term is lying in the direction of the choice of optimal configurations for interacting waves, i.e. in the direction of improving DIA. Several variants of improvements of DIA will be considered in the following section.

6 Examples of improved DIA

As can be seen from the estimations in Sect. 5, there are essentially two ways for improving DIA: (1) to enhance the speed (and accuracy) of calculation for one-configuration DIA, or (2) to enhance an accuracy of approximation by adding new configurations (multiple DIA). Here we shall give two examples of improved DIA using these ways.

6.1 Fast DIA adjusted to the integration grid

As was mentioned in Sect. 2.1, the technical shortage of the original (one-configuration) DIA used in WAM is the presence of intermediate and cumbersome interpolation procedures provided by a mismatch of the integration grid and the vectors \mathbf{k}_+ , \mathbf{k}_- , which leads increase of computation time. Thus, to enhance a speed of calculation in a one-configuration DIA, one could adjust the configuration of interacting waves to the integration grid, and thus, avoiding interpolation. It can be done in the following way.

Instead of configuration (10) we propose to use the following:

- *k*₄ = *k*, where the arbitrary wave vector *k* is located at a grid node and represented in the polar co-ordinates by σ and θ;
- 2. $k_3 = k_+$ where k_+ is represented by

$$\sigma_3 = \sigma(1 + \alpha_{34}) \text{ and } \theta_3 = \theta + \Delta \theta_{34};$$
 (44)

3. $\mathbf{k}_1 \approx \mathbf{k}_2 \approx (\mathbf{k}_4 + \mathbf{k}_3)/2 \equiv \mathbf{k}_a$ where \mathbf{k}_a is represented by

$$\theta_a = \theta + \Delta \theta_{a4}$$

The choice of the latter two vectors depends on the values of the parameters σ_{34} and $\Delta \theta_{34}$ defining vectors k_3 and k_a . To solve the problem posed, one needs to meet the following requirements: (a) to preserve approximately the original DIA configuration, (b) to allocate all the vectors at the grid nodes, and (c) to chose values of α_{34} and $\Delta \theta_{34}$ in such a manner that the interacting vectors are allocated at the "figure-of eight" in the *k*-space (see Fig. 1).

The first requirement (original DIA configuration) can be expressed numerically as follows

$$\alpha_{34} \approx 1.5, \ \sigma_1 \approx \sigma_2, \ \text{and} \ \Delta \theta_{34} = 45^{\circ}.$$
 (45)

The second requirement (allocation of vectors at the grid nodes) is expressed by the following Fortran statements

$$Mod(\theta_{34}, \ \Delta\theta_g) \ll \Delta\theta_g, \tag{46}$$
$$Mod(\theta_{a4}, \ \Delta\theta_g) \ll \Delta\theta_g,$$



Fig. 5. Comparative diagram of relative errors for the fast and original DIA. For legend see Fig. 1.

and by the mathematical expressions

$$\sigma_3/\sigma \approx e^{m3}, \, \sigma_1/\sigma \approx e^{m1}, \, \sigma_2/\sigma \approx e^{m2}.$$
 (47)

In Eqs. (46)and (47), *e* is the frequency grid exponent, $\Delta \theta_g$ is the angle grid resolution (see Eqs. 35 and 36), and powers *m*1, *m*2, *m*3 are integer values to be found.

The third requirements (allocation at the "figure-of eight") is expressed by the ratio (Polnikov, 1989):

$$k_a = \sigma_a^2 / 2,\tag{48}$$

where

$$k_a = \left[\sigma^4 + \sigma_3^4 + 2\sigma^2 \sigma_3^2 \cos(\Delta \theta_{34})\right]^{1/2}$$
(49)

and

$$\sigma_a = \sigma + \sigma_3. \tag{50}$$

The expression for $\Delta \theta_{34}$ is deduced from the resonant condition (5),

$$\Delta \theta_{a4} = \arctan\left[\frac{\sigma_3^2 \sin(\Delta \theta_{34})}{\sigma_3^2 \cos(\Delta \theta_{34}) + \sigma^2}\right].$$
(51)

To meet all three requirements simultaneously, it is sufficient to fit properly a value of m3 in Eq. (47) for the fixed grid parameters e and $\Delta \theta_g$. In the case of a reasonable grid resolution, the proper values of m1, m2, m3, $\Delta \theta_{a4}$, and $\Delta \theta_{34}$ can be found without any difficulties. In particular, for the grid given by Eqs. (35) and (36), we have

$$m3 = 10, ml = 5, m2 = 6;$$

$$Int(\Delta\theta_{34}/\Delta\theta_g) = 4, Int(\Delta\theta_{a4}/\Delta\theta_g) = 3$$
(52)

where the function Int(...) means the nearest integer number to the argument. Equation (52) represent the fast DIA for the WAM-configuration adjusted to the grid (35), (36). The numerical integration part is absolutely the same as for original DIA and described by equations of the type (11) and (12).

Table 5. MRE for NL-transfer in fast DIA with WAM-configuration

No of run	1	2	3	4	5	6	7	15	16
Coeff. C_{ad}	1.99	1.56	2.11	2.83	1.57	1.48	1.89	1.90	3.2
MRE(10%)	28.9	30.8	67.6	62.1	29.5	40.1	48.6	65.7	63.5
MRE(20%)	20.9	20.3	66.5	51.2	24.8	30.8	29.0	63.5	39.7
MRE(10%)	26.3	17.5	66.3	56.4	75.6	74.8	56.7	80.9	100
For 1D-NL									

Final results of the mean relative errors for the representative set of spectral parameters are presented in Table 5 and Fig. 5.

From Fig. 5 it is seen that the fast one-configuration DIA has better features with respect to original DIA. In essence, we mention only that the expert MRE in this case is

$$\varepsilon_{rel}^{(fast\ DIA)} \cong 0.48$$
(53)

whilst the relative speed has the maximum possible value

$$t_{rel}^{(fast\ DIA)} = T_{ref}/\tau = 1.$$
(54)

From Eqs. (54) and (24), (25) it follows that one should use the first type of efficiency. Finally, the efficiency of the approximation under consideration is of the order

$$Eff_1 = (1/0.48)^2 \cong 4.3.$$
 (55)

Thus, the proposed fast DIA with WAM-configuration has twice the efficiency of original DIA¹².

6.2 Modified (multiple) DIA adjusted to integration grid

As follows from above, the advantage of the fast WAMconfiguration DIA is provided by the enhancement of speed of calculation only. To enhance the accuracy of approximation, one should construct DIA with several configuration of interacting waves (the so-called multiple DIA, Van Vledder et al, 2000). Here we represent one of such a multiple DIA with three configurations adjusted to the integration grid given by Eqs. (35) and (36) (3c-DIA).

The main idea of a multiple configuration DIA is to involve into the approximation of *NL*-term such configurations, which cover an interaction grid wider in frequency and angle space than it is done in the WAM-configuration. The simplest way is to vary the parameters α_{34} and $\Delta \theta_{34}$ in Eq. (44). Herewith, one should take in mind the weights of the configurations, B_c , as it follows from Eq. (18). But this point can be easily done by a previous estimation of B_c from the exact integrand expressions presented in integral (17).¹³

By analogy with the procedure described in Eq. (44), in Sect. 6.1, we have found the following version of the 3c-DIA:



Fig. 6. Comparative diagram of relative errors for the 3c-DIA and fast DIA. Only the expert set of runs from Table 1 is shown.

1. Configuration 1 is presented by ratios

$$m3 = 8, m1 = 4, m2 = 5;$$

$$Int(\Delta\theta_{34}/\Delta\theta_g) = 3, Int(\Delta\theta_{a4}/\Delta\theta_g) = 2;$$
 (56)

2. Configuration 2 is given by Eq. (52)

$$m3 = 12, m1 = 5, m2 = 6;$$

 $Int(\Delta\theta_{34}/\Delta\theta_g) = 4, Int(\Delta\theta_{a4}/\Delta\theta_g) = 3;$ (57)

3. Configuration 3 is presented by ratios

$$m3 = 12, m1 = 7, m2 = 7;$$

 $Int(\Delta\theta_{34}/\Delta\theta_g) = 5, Int(\Delta\theta_{a4}/\Delta\theta_g) = 4;$ (58)

For these configurations the relative values of weights, Bc, are approximately equal to each other. The rest part of calculations in 3c-DIA is nearly the same as in the fast DIA (exception is the number of configurations). Results of testing the three-configuration DIA, given by Eqs. (56)–(58), are presented in Table 6 and Fig. 6 for the representative set of spectral parameters.

Comparing of results shows that the 3c-DIA has better features than the fast DIA (Fig. 6) and much better features with

¹²An implementation of fast DIA into the WAM (Cycle 4) gives the value of RP about 25%. In details the point of modified DIA implementation will be discussed in another paper.

¹³From this point of view, the multiple DIA resembles RIA. But the difference is in the choice of configurations.

Table 6. MRE for NL-transfer in three-configuration DIA

No of run	1	2	3	4	5	6	7	15	16
Coefficient C_{ad}	0.78	0.63	1.06	1.07	0.62	0.57	0.73	0.85	1.07
MRE(10%)	25.2	28.8	53.8	37.6	28.9	37.4	41.6	49.9	52.6
MRE(20%)	17.1	19.5	49.8	40.3	24.9	28.3	26.2	45.2	37.3
MRE(10%)	18.6	20.9	62.8	52.6	62.4	65.9	55.4	69.9	85.1
for 1D-NL									

respect to the original DIA (compare Figs. 5 and 6). In this case, the expert MRE is

$$\varepsilon_{rel}^{(3c-DIA)} \cong 0.39\tag{59}$$

and the relative speed has the value

$$t_{rel}^{(3c-DIA)} = T_{ref}/\tau = 0.33.$$
 (60)

From Eq. (60) it follows that one should use the second type of efficiency. By analogy with the said above in Eq. (37), above in Sect. 6.1, we may take a relative part RP of the order of 1/3. Finally, we find

$$Eff_1 \cong (1/0.39)^2 (0.33)^{1/3} \cong 4.4.$$
 (61)

Comparing Eq. (55) and Eq. (61) gives that the proposed three-configuration DIA has nearly the same efficiency as the fast DIA with the WAM-configuration. Herewith, taking in mind the better accuracy, we may state that the former approximation is preferable.

7 Conclusions

The problem of optimizing approximations of the kinetic integral describing the rate of four- wave nonlinear interactions is not completely solved yet. However, at present it seems that this problem is very close to its solution. The ground for this statement is the finding of this work that all the most prospective approximations: Discrete Interaction Approximation (DIA), Reduced Integration Approximation (RIA), and Diffusion Approximation (DA), have the same root. All of them are different modifications of the Discrete Interaction Approximation. The only difference among them is the choice of configurations for the interacting waves. Thus, the solution of the problem is the proper choice of the interacting configurations in a modified DIA.

To make this choice, one needs a tool for estimation of accuracy and efficiency of an approximation. This tool includes formal procedures on how to estimate these two parameters. On the basis of several heuristic postulates, appropriate formulas were constructed in this paper. They are Eqs. (24) and (25) for the efficiency parameter, Eff, and Eq. (31) for the mean relative error of approximation, ε_{rel} .

The formulas mentioned above were used for comparative estimation of accuracy and efficiency of two approximations: original DIA (WAM-configuration) and DA in Zakharov-Pushkarev's version (ZPA). It was shown that despite the smaller computational time in ZPA, the accuracy and efficiency of DIA is higher.

Taking in mind that the technical shortage in the original DIA (Hasselmann et al., 1985) consists in the presence of an interpolation procedure aimed to solve a mismatch of the integration grid nodes and interacting wave vectors, two modification of DIA were proposed. The first is the so-called "the fast one-configuration DIA" (fast DIA), the second is the three-configuration DIA (3c-DIA). Both of them are adjusted to the certain integration grid, to exclude the cumbersome interpolation procedure mentioned above. Owing to this feature, both new approximations have efficiency more than twice greater than the original DIA. Herewith, the fast DIA is quicker in calculation, but the 3c-DIA has more accuracy. Both of approximations proposed are feasible to be implemented into practice, and this point will be described in another paper.

Here we did not test RIA given by expression (16). Nevertheless, one may state in advance, that due to similarity RIA to a multiple configuration version of DA, efficiency of RIA is expected to be smaller than one for the original DIA. This point may be verified later, if necessary.

Due to the fact that the found values of accuracy and efficiency for the proposed modifications of DIA are far from their upper limits provided by the aim of practice, more effective approximations could be constructed, in principle. This is the final task of the problem under consideration.

Acknowledgements. The authors are grateful to the administration of CPTEC and CNPq for the funding this work. In part, this work was supported by the Russian Fund for Basic Research, project # 01-05-64580. We appreciate Gerbrant van Vledder for his remarks made during the reading of the proofs.

References

- Barnett, T. P.: On the generation, dissipation and prediction of ocean wind waves, J. Geophys. Res., 73, 513–529, 1968.
- Efimov, V. and Polnikov, V.: Numerical modeling of wind waves, Kiev, Naukova dumka Publishing House, 240, 1991.
- Crawford, D. R., Saffman, P. G., and Yuen, H. C.: Evolution of a random inhomogeneous field of nonlinear deep-water gravity waves, J. Wave Motion, 2, 1–16, 1980.

V. G. Polnikov and L. Farina: The four-wave kinetic integral

- Hashimoto, N.: Extension of the DIA for computing the nonlinear energy transfer of ocean gravity waves, Proc. Jap. Conf. On Coastal Egnineering, (in Japanese), 1999.
- Hashimoto, N. and Kawagushi, K.: Extension and modification of the Discrete Interaction Approximation (DIA) for computing nonlinear energy transfer of gravity wave spectra, Proc. 4th Int. Symp. on Ocean Waves, Measurement and Analysis, WAVES-2001, 2001.
- Hasselmann, K.: On the non-linear energy transfer in a gravity wave spectrum, Pt. 1. General theory, J. Fluid Mech., 12, 481–500, 1962.
- Hasselmann, S. and Hasselmann, K.: A symmetrical method of computing the nonlinear transfer in a gravity wave spectrum, Hamburger Geophys. Einzelschrift, 52, 138, 1981.
- Hasselmann, S., Hasselmann, K., Allender, K. J. and Barnett, T. P.: Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part II, J. Phys. Oceanogr., 15, 1378–1391, 1985.
- Jenkins, A. and Phillips, O. M.: A Simple Formula for nonlinear Wave-Wave Interaction, Intern. J. of Offshore and Polar Engineering, 11, 81–86, 2001.
- Komen, G. J., Cavaleri, L., Donelan, M., et al.: Dynamics and Modeling of Ocean Waves, N.Y., Cambridge University Press, 532, 1994.
- Lavrenov, I. V. and Polnikov, V. G.: A study of properties for non-stationary solutions of the Hasselmann's kinetic equation, Izvestiya, Atmos. Oceanic. Phys., (English transl.), 37, 5, 2001.
- Lin, R. Q. and Perry, W.: Wave-wave interactions in finite depth water, J. Geophys. Res., 104, 11 193–11 213, 1999.
- Masuda, A.: Nonlinear energy transfer between wind waves, J. Phys. Oceanogr., 15, 1369–1377, 1980.
- Polnikov, V. G.: Calculation of the nonlinear energy transfer through the surface gravity waves spectrum. Izv. Acad. Sci. SSSR, Atmos. Oceanic. Phys., (English transl.), 25, 896–904, 1989.
- Polnikov, V. G.: Numerical solution of the kinetic equation for surface gravity waves, Ibid., (English transl.), 26, 118–123, 1990.
- Polnikov, V. G.: A third generation spectral model for wind waves,

Ibid., (English transl.), 27, 615–623, 1991.

- Polnikov, V. G.: Numerical modeling of flux spectra formation for surface gravity waves, J. Fluid Mech., 278, 289–296, 1994.
- Polnikov, V. G.: Numerical modeling of the constant flux spectra for surface gravity waves in a case of angular anisotropy, Wave Motion, 33, 271–282, 2001.
- Polnikov, V. G.: A Basing of the Diffusion Approximation Derivation for the Four-Wave Kinetic Integral and Properties of the Approximation, Nonl. Proc. Geophys., 9, 3/4, 355–366, 2002.
- Resio, D. T.: The estimation of wind wave generation in a discrete spectral model, J. Phys. Oceanogr., 11, 510–525, 1981.
- Resio, D. and Perry, W.: A numerical study of nonlinear energy fluxes due to wave-wave interactions., J. Fluid Mech., 223, 603– 629, 1991.
- The SWAMP group: Ocean wave modeling, N. Y. and L., Plenum press, 256, 1985.
- WAMDI group: The WAM model A third generation Ocean Wave Prediction Model, J. Phys. Oceanogr., 18, 1775–1810, 1988.
- Webb, D. J.: Non-linear transfer between sea waves, Deep Sea Res., 25, 279–298, 1978.
- Zakharov, V. E. and Filonenko, N. N.: The energy spectrum for stochastic oscillation of fluid's surface, Dokl. Akad. Nauk., 170, 1292–1295, 1966.
- Zakharov, V. E. and Smilga, A. V.: On quasi-one-dimensional spectra of weak turbulence, Sov. Phys. JETP, (English transl.), 54, 701–710, 1981.
- Zakharov, V. E. and Pushkarev, A.: Diffusion model of interacting gravity waves on the surface of deep fluid, Nonl. Proc. Geophys., 6, 1–10, 1999.
- Van Vledder, G. Ph., Herbers, T. H. C., Jense, R. E., et al.: Modeling of nonlinear quadruplet wave-wave interactions in operational model, Proc. 27th Int. Conf. Coastal Engineering, 2000.
- Van Vledder, G. Ph.: Extension of the Discrete Interaction Approximation for computing nonlinear quadruplet wave-wave interactions in operational wave prediction model, Proc. 4th Int. Conf. on Ocean Waves. WAVES-2001. San Fransisco, Sept., 2001.
- Young, I. R. and Van Vledder, G. Ph.: The Central Role of Nonlinear Interactions in Wind-wave Evolution, Phil. Trans. Roy. Soc. Lond., A 342, 505–524, 1993.